

**UNIVERSIDAD COMPLUTENSE DE MADRID**  
**FACULTAD DE MEDICINA**  
**Departamento de Microbiología I**



**TESIS DOCTORAL**

**Estudio del polimorfismo del gen inmunosupresor HLA-G, en  
amerindios**

**MEMORIA PARA OPTAR AL GRADO DE DOCTOR**

**PRESENTADA POR**

**Mercedes Enríquez de Salamanca Carrascosa**

**Director**

**Antonio Arnaiz Villena**

**Madrid, 2018**

**UNIVERSIDAD COMPLUTENSE DE MADRID**

**FACULTAD DE MEDICINA**

**DPTO. MICROBIOLOGÍA I**



**ESTUDIO DEL POLIMORFISMO DEL GEN  
INMUNOSUPRESOR HLA-G, EN  
AMERINDIOS.**

**MERCEDES ENRÍQUEZ DE SALAMANCA  
CARRASCOSA**

**DIRECTOR: ANTONIO ARNAIZ VILLENA**

**MADRID, 2017**

Diseño de Portada: Gerónimo, Jefe Apache (Arriba Izqda.); Mapa de América (Arriba Dcha.), los colores representan la hipótesis de Greenberg de poblamiento de América; Sistema HLA (Abajo Izqda); Kipu (Abajo Dcha.), utilizado para comunicarse por los amerindios andinos. (Por José Palacio-Gruber y Antonio Arnaiz-Villena)

# **TESIS DOCTORAL**

**TÍTULO: ESTUDIO DEL POLIMORFISMO DEL GEN  
INMUNOSUPRESOR HLA-G, EN AMERINDIOS.**

**AUTORA: MERCEDES ENRÍQUEZ DE SALAMANCA  
CARRASCOSA**

**DIRECTOR: PROF. DR. ANTONIO ARNAIZ VILLENA**

Catedrático de Inmunología y Coordinador I+D  
Centro de Transfusión de la Comunidad de Madrid  
Universidad Complutense de Madrid

**LUGAR DE REALIZACIÓN:**

- **Departamento de I+D del Centro de Transfusión de la Comunidad de Madrid**
- **Departamento de Microbiología I (Inmunología) de la Facultad de Medicina de la Universidad Complutense de Madrid**



**Don Antonio Arnaiz Villena, Doctor en Medicina y Cirugía, Catedrático de Inmunología de la Universidad Complutense de Madrid y Coordinador de I+D del Centro de Transfusión de la Comunidad de Madrid,**

**CERTIFICA:**

**Que Doña Mercedes Enríquez de Salamanca Carrascosa, ha realizado bajo su dirección el trabajo de Tesis Doctoral que lleva por título:**

**“Estudio del polimorfismo del gen inmunosupresor HLA-G, en amerindios”.**

**Revisado el presente trabajo, considera que tiene la debida calidad para su defensa y calificación.**

**Fdo.: Prof. Dr. Antonio Arnaiz Villena**



# **AGRADECIMIENTOS.**

La correcta realización de este trabajo, no hubiese sido posible sin la ayuda y colaboración de ciertas personas que, en mayor o menor medida, han guiado mi trabajo.

Por ello, me gustaría agradecer, en primer lugar, a mi director el Prof. Dr. Antonio Arnaiz Villena, por haberme brindado la oportunidad de trabajar en su equipo durante estos años y por haberme aportado todo el conocimiento, apoyo e indicaciones necesarias para realizar esta tesis doctoral.

A mis compañeros, Carlos, Diego y Cristina, por haberme enseñado todas las técnicas imprescindibles para la realización de este proyecto, por su paciencia y comprensión. En especial quiero agradecer a Ester su ayuda en los últimos momentos de esta experiencia, por correcciones y trámites burocráticos de los que se ha encargado.

Al Ministerio de Sanidad y a la Fundación Mutua Madrileña, por haber apoyado a esta investigación, con los fondos precisos.

Al Centro de Transfusión de la Comunidad de Madrid, por haberme ofrecido realizar en sus instalaciones, mi trabajo y a las personas que trabajan en él, por su amabilidad durante estos años.

A la Universidad Complutense de Madrid y al Departamento de Inmunología de la Facultad de Medicina, por haber realizado en él mis estudios de doctorado y haber logrado así obtener esta Tesis Doctoral, especialmente a su secretaria Lola por su apoyo necesario.





# DEDICATORIA

Este trabajo se lo quiero dedicar, en primer lugar y por encima de todo, a Santi. Por estar siempre a mi lado, pendiente de los avances del trabajo, por darme consejo en todo momento, por animarme incansablemente a seguir adelante a pesar de las dificultades, sin perder en ningún momento la sonrisa ni la paciencia, por facilitarme siempre el trabajo con su alegría, por su ejemplo de responsabilidad y trabajo duro, en definitiva, por ser mi pilar de apoyo.

A Macarena, Jacobo y Javier, por ser mi motivo de alegría en todos los aspectos y en particular, en esta dura meta personal, por haber sufrido conmigo, de una forma tan estrecha, estos últimos años de trabajo y por ser parte de mí.

Sin ellos cuatro, este reto nunca podría haber llegado a su fin.

A mi gran familia, en especial a mis padres, por animarme bajo cualquier circunstancia, ayudarme a iniciar este camino, por ser mi ejemplo de vida y porque sin ellos no sería quien soy y a mis hermanos, abuela, ahijada, sobrino y cuñado, por haber sabido sufrir con paciencia mis malos momentos.

A mis amigos, por ser una fuente de inspiración para mí, por mostrarme con sus ejemplos que merece la pena la lucha para conseguir lo que uno desea y por compartir conmigo la pasión por el mundo de la ciencia.

A mi tío, por haberme animado a comenzar esta aventura que sin él no habría logrado concluir y por apoyar, sin descanso, a toda la familia siempre que se le ha necesitado.

# ÍNDICE



SUMMARY.....	1
RESUMEN.....	8
INTRODUCCIÓN.....	15
I. GENERALIDADES DEL SISTEMA HLA.....	17
1. DEFINICIÓN.....	17
2. HISTORIA.....	18
3. LOCALIZACIÓN GENÓMICA Y CLASIFICACIÓN.....	19
3.1 Localización.....	20
3.2 Clasificación.....	20
3.2.1 Región de clase I extendida.....	21
3.2.2 Región de clase I.....	21
3.2.3 Región de clase II.....	22
3.2.4 Región de clase II extendida.....	22
3.2.5 Región de clase III.....	22
3.2.6 Repeticiones intercaladas.....	23
4. ESTRUCTURA DEL SISTEMA HLA.....	23
4.1 Sistema HLA de clase I.....	23
4.1.1 Estructura génica de los antígenos HLA de clase I, clásicos.....	23
4.1.2 Estructura proteica de los antígenos HLA de clase I, clásicos.....	24
4.1.3 Estructura génica de los antígenos HLA de clase I, no clásicos.....	28
4.1.4 Estructura proteica de los antígenos HLA de clase I, no clásicos....	28
4.2 Sistema HLA de clase II.....	30
4.2.1 Estructura génica de los antígenos HLA de clase II.....	30
4.2.2 Estructura proteica de los antígenos HLA de clase II.....	31
5. NOMENCLATURA.....	34
6. FUNCIÓN DE LAS MOLÉCULAS HLA: PRESENTACIÓN ANTIGÉNICA.....	36
6.1 Moléculas HLA de clase I.....	37
6.2 Moléculas HLA de clase II.....	38
7. HLA Y FARMACOGENÓMICA.....	40
8. HLA Y TRASPLANTE.....	43
9. HLA Y ENFERMEDAD.....	43

II. HLA-G: MOLÉCULA Y GEN.....	45
1. ESTRUCTURA.....	45
1.1 Estructura génica.....	45
1.2 Estructura proteica.....	46
1.3 Polimerización.....	49
2. POLIMORFISMO.....	50
2.1 Polimorfismo genético.....	50
3. FUNCIÓN.....	54
3.1 Modulación de la respuesta inmune.....	54
3.2 HLA-G y su importancia en la gestación.....	55
3.3 HLA-G y su papel en trasplante.....	57
3.4 HLA-G y su doble función en cáncer.....	58
3.5 HLA-G y su relación con enfermedades autoinmunes.....	59
3.6 HLA-G y su relación con el VIH.....	60
III. POBLACIÓN DE ESTUDIO: AMERINDIOS INMIGRANTES EN MADRID...	63
1. EL ORIGEN DE LOS AMERINDIOS.....	63
2. ESTUDIO DEL SISTEMA HLA EN POBLACIONES AMERINDIAS.....	64
2.1 Polimorfismos de los genes HLA en amerindios.....	64
2.2 Relaciones genéticas en las poblaciones amerindias.....	66
3. AMERINDIOS INMIGRANTES EN MADRID.....	67
<b>OBJETIVOS.....</b>	<b>70</b>
<b>MATERIAL Y MÉTODOS.....</b>	<b>73</b>
1. SELECCIÓN DE MUESTRAS.....	75
2. OBTENCIÓN DE MUESTRAS.....	75
3. PREPARACIÓN DE LAS MUESTRAS.....	75
4. EXTRACCIÓN DE ADN GENÓMICO.....	75
5. AMPLIFICACIÓN DE SECUENCIAS HLA-G.....	76
5.1 Control de la amplificación.....	79
5.2 Purificación de las bandas.....	80
6. SECUENCIACIÓN DE LOS FRAGMENTOS AMPLIFICADOS.....	81

7. DETERMINACIÓN DE ALELOS HLA-G.....	83
8. ANÁLISIS ESTADÍSTICOS.....	84
8.1 Cálculo de frecuencias alélicas.....	84
8.2 Estimación de los haplotipos extendidos y desequilibrio de ligamiento.....	84
8.3 Test de Fisher.....	85
9. CARACTERIZACIÓN DE LOS NUEVOS ALELOS.....	86
9.1 Repetición de la secuenciación directa.....	86
9.2 Repetición de la determinación de los alelos.....	86
9.3 Secuenciación mediante clonaje del producto amplificado.....	87
9.4 Obtención de las secuencias de ADN de los fragmentos clonados.....	89
9.5 Depósito de los nuevos alelos de HLA-G .....	90
9.6 Determinación de las distancias genéticas.....	90
9.7 Determinación de las secuencias aminoacídicas de los alelos nuevos.....	91

## RESULTADOS..... 92

1. DETERMINACIÓN DE LOS ALELOS HLA-G EN LA POBLACIÓN AMERINDIA INMIGRANTE EN MADRID.....	94
2. CÁLCULO DE LAS FRECUENCIAS ALÉLICAS HLA-G EN LA POBLACIÓN AMERINDIA INMIGRANTE EN MADRID.....	96
3. ANÁLISIS DE LOS HAPLOTIPOS EXTENDIDOS EN LA POBLACIÓN AMERINDIA INMIGRANTE EN MADRID.....	99
3.1 Loci HLA-A y -G.....	99
3.2 Loci HLA-B y -G.....	101
3.3 Haplotipo extendido HLA-A-B-G.....	102
3.4 Haplotipo extendido HLA-A-DRB1-G.....	103
3.5 Haplotipo extendido HLA-A-B-DRB1-G.....	104
4. ESTUDIO DEL ALELO NULO, HLA-G*01:05N, EN LA POBLACIÓN AMERINDIA INMIGRANTE EN MADRID.....	106
4.1 Descripción alelos nulos encontrados en la muestra.....	106
4.2 Frecuencias de los alelos nulos encontrados en la muestra.....	107
4.3 Alelo nulo y los haplotipos extendidos.....	107
4.4 Comparación de las frecuencias del alelo nulo de la población de estudio con otras poblaciones mundiales estudiadas.....	108

4.5 Alelo nulo y el desequilibrio de ligamiento con el locus HLA-A.....	111
5. ESTUDIO DE LAS PROTEÍNAS HLA-G ENCONTRADAS EN LA POBLACIÓN AMERINDIA INMIGRANTE EN MADRID.....	111
5.1 Secuencias y estructuras de los nuevos alelos. ....	111
5.1.1 Un alelo nuevo en los Andes, Ecuador .....	111
5.1.2 Un alelo nuevo en el Caribe, República Dominicana.....	117
5.2 Distancias genéticas.....	123

## DISCUSIÓN..... 126

1. HLA-G Y SU POLIMORFISMO EN LA POBLACIÓN AMERINDIA INMIGRANTE EN MADRID.....	128
2. ANÁLISIS DE LOS HAPLOTIPOS EXTENDIDOS.....	130
3. ANÁLISIS DEL ALELO NULO, HLA-G*01:05N, EN LA POBLACIÓN AMERINDIA.....	132
3.1 Alelo nulo en Uros y Mayas.....	133
3.2 Alelo nulo en amerindios inmigrantes en Madrid.....	134
4. ALELOS NUEVOS Y SUS POSIBLES IMPLICACIONES A NIVEL MOLECULAR.....	137
5. LA FISIOLOGÍA Y PATOLOGÍA DE HLA-G.....	141

## CONCLUSIONES..... 143

## BIBLIOGRAFÍA..... 146

## ANEXOS..... 160

ANEXO I: CUESTIONARIO PARA DONANTES VOLUNTARIOS.....	162
ANEXO II: SIMBOLOGÍA EN EL ADN Y CÓDIGO GENÉTICO.....	164
ANEXO III: INFORMACIÓN SOBRE AMINOÁCIDOS.....	165
ANEXO IV: LISTADO DE ALELOS HLA-G EN LA ACTUALIDAD.....	166
ANEXO V: ABREVIATURAS MÁS FRECUENTES EN EL TEXTO.....	167
ANEXO VI: LISTADO DE ALELOS HLA-G EN LA POBLACIÓN DE ESTUDIO .....	169



ANEXO VII: CURRICULUM VITAE.....	173
ANEXO VIII: TRABAJOS PUBLICADOS.....	180

# SUMMARY



## **INTRODUCTION.**

The Major Histocompatibility Complex (MHC) is a set of cell surface glycoproteins molecules encoded by a large gene family which controls a major part of the immune system in all vertebrates. The MHC determines compatibility of donors for organ transplant, as well as one's susceptibility to an autoimmune disease via cross reacting immunization. The Human Leukocyte Antigen (HLA) is the name for the human MHC.

The HLA is an extended region of the genome that spans some 4 million base pairs (bp) on the short arm of human chromosome 6 between 6p21.3. The major function of the HLA molecules is to facilitate the display of unique molecular fragments on the surface of cells in an arrangement that permits their recognition by immune effectors such as T-lymphocytes. The HLA-I or HLA-II molecules are those cell surface glycoproteins that actually perform the binding and recognition steps, while other genes that map to these regions may contribute to antigen-processing and presentation functions in other distinct ways.

A polymorphic locus or gene is one that has a high frequency of genetic variants and the HLA genes are the most polymorphic genes in man. This is understandable in light of the central role these play in the immune response.

HLA-I and HLA-II are not only two distinct genetic regions within the HLA system, with the HLA-III region interspersed between them, but also code for structurally different proteins. HLA-I molecules consist of a heavy chain ( $\alpha$  chain), which is non-covalently assembled with a light chain known as  $\beta$ 2-microglobulin ( $\beta$ 2-m, encoded by the b2m gene on chromosome 15). HLA-II molecules are also heterodimers, composed of an  $\alpha$  chain noncovalently assembled with a  $\beta$ 2-m chain. Both HLA-I and HLA-II proteins have similar structure consisting of extra cellular region, trans-membrane region and a short intra-cytoplasmic tail. Class-I protein has three domains ( $\alpha$ 1,  $\alpha$ 2 and  $\alpha$ 3) of which the membrane distal  $\alpha$ 1 and  $\alpha$ 2 domains form the antigen-binding groove. The  $\alpha$ 3 domain along with  $\beta$ 2-m form the membrane proximal region. The  $\alpha$  and  $\beta$  chains of class-II have two domains each and the membrane distal  $\alpha$ 1 and  $\beta$ 1 domains form the peptide groove while  $\alpha$ 2 and  $\beta$ 2-m are membrane proximal regions.

HLA-G is a class I, non-classical molecule, whose structure resembles other classic class I HLA molecules. The HLA-G gene has low polymorphism in its coding region (only 53 alleles are described to date), a limited expression pattern in healthy conditions, and the unique characteristic among HLA molecules of forming multimers. In addition, seven different isoforms can be formed through alternative splicing (HLA-

G1 to HLA-G7), four membrane-bound, G1-G4, and three soluble, G5-G7. All those characteristics contribute for the increasing scientific interest on that molecule, some of which have vital importance by the induction of immune tolerance. In the HLA-G molecule, the peptide is located more deeply into the cleft between domains  $\alpha 1$  and  $\alpha 2$  as compared with classic HLA molecules. These special characteristics of HLA-G make it unlikely that this molecule plays an important role in antigen presentation. The expression of HLA-G is highly restricted to specific tissues, in addition to being expressed in fetal tissues, such as the trophoblast cells, it is constitutively expressed in adult thymus, cornea, pancreatic islets, and precursors of endothelial and erythroid cells. However, HLA-G expression can be induced in situations such as transplantations, inflammatory diseases, tumour cells, multiple sclerosis, and viral infections.

The peopling of America, and the source of Native American populations, has been the subject of various studies. There are several theories about the peopling of America. One of the most significant is the Greenberg's theory. He first postulated the triple migrations theory; Amerindians (most North and South American Indians; between 30.000- 15.000 years before the present (BP)), Na-Dene (Athabascans, Navajo, Apache; between 15.000-10.000 BP) and Eskimo-Aleuts (10.000-8.000 BP). Secondly, a trans-Pacific route of American peopling from Asia or Polynesia has been suggested, but finally, both genetic and archaeological evidence suggests that a two-ways trans-Atlantic traffic occurred before Columbus discovered America; archaeologists in New Mexico have recently found tools used 20,000 years ago, in Spain.

Most Amerindian immigrants in Madrid come from Andean Countries in the last 10 years (mainly Ecuador, Colombia, Peru and Bolivia). They show an HLA profile with "quasi-specific alleles", which makes them different to the rest of the World, and they never had been typed for HLA-G gene before. Immigrant volunteer unrelated blood donors contributed to the present study.

### **AIMS.**

The aims of this work are as follows:

- 1- To determine HLA-G alleles in a population of immigrant Amerindians in Madrid.
- 2- Calculate the extended haplotypes HLA-G / HLA-A / HLA-DRB1 (class I and class II) in the same population.
- 3- To compare the frequencies of Amerindian HLA-G alleles with those of the rest of the world populations.

- 4- Search for possible new alleles of the HLA-G locus in Amerindians.
- 5- To extract functional and evolutive conclusions about HLA-G from the results (mainly of the null allele HLA-G \* 01: 05N), in the different populations of the world.

## RESULTS.

A total of one hundred fifty-four individuals, have certain HLA-G alleles with a four-digit resolution, and one hundred seventeen individuals were determined with six digits of resolution. In both cases individuals are Amerindian immigrants in Madrid, healthy and unrelated.

Of all 53 alleles described, only eighteen are represented in this population. The most common groups are; HLA-G\*01:01 (71.8%), HLA-G\*01:04 (18.5%), HLA-G\*01:03 (5.8%) and HLA-G\*01:06 (2.9%).

The most common extended haplotype observed in this population is HLA-A\*24:02-B\*35:01-DRB1\*14:02-G\*01:04, with a frequency bordering on 3% of the population. This haplotype, without the locus HLA-G, is also found in other Amerindian populations such as Aymara (2.8%), Quechua (1.4%) and Jaidukama (7.7%).

It should be mentioned that, in the rest of world populations, extended haplotypes with HLA-G locus have not been studied, so when compare haplotypes with other populations, no one considers this locus.

The second and third most frequent extended haplotype HLA-A\*02:01-B\*35:04-DRB1\*08:02-G\*01:01 (1.7%) and HLA-A\*02:01-B\*40:02-DRB1\*04:07-G\*01:01 (1.2%) respectively, can also be seen in numerous Amerindian populations. Furthermore, there is a certain European influence, which can be observed in HLA-A\*03:01-B\*07:02-DRB1\*15:01-G\*01:01, since it has a possible North African/Western Europe origin and has a high frequency in populations of, for example, the North of Spain and Ireland.

There are three null alleles, HLA-G\*01:05N, found in three different heterozygous individuals. The presence of this allele, was also observed with a low frequency (1-1.3%, four and six-digit resolution respectively). This also applies to European Caucasian populations, Oriental, Pygmies in African and Brazilian. In the case of Amerindian populations, Mayans and Uros, the presence of this allele is null. Three extended haplotypes carrying the null allele in this population, HLA-A\*33:03-B\*57:03-DRB1\*13:02-G\*01:05N, HLA-A\*31:01-B\*15:05-DRB1\*04:03-G\*01:05N and HLA-

A\*29:02-B\*15:05-DRB1\*04:03-G\*01:05N, all of them with the lowest frequency encountered (0,6%) and none have been observed in other populations. Only significant differences were observed ( $p < 0.05$ ) among immigrants Amerindians in Madrid and the Middle East populations (North Indians, Iranians and Iraqi) and African Negroids, except Pygmies (Africans from Ghana, African American and African Shona). The latest have the highest frequencies of this allele. Moreover, corroborating the above in previous studies, the null allele HLA-G\*01:05N is in linkage disequilibrium with HLA-A\*30:01 also in Amerindian immigrant of Madrid's population.

During the HLA-G alleles typing process, we found two new alleles. Two heterozygous individuals of different origin: The first, of Ecuadorian origin (new allele 1) and the second (new allele 2), of Dominican origin. In the case of the new allele 1 they were determined exons 2, 3 and 4. Throughout these three exons, four positions can be located where the nucleotide changes, creating an unseen new sequence in any individual above.

Regarding new allele 2, were determined exons 2 and 3. The nucleotide localized variations in this allele were only found in exon 3.

Thus, three Single Nucleotid Polymorphism (SNPs) were observed, of which the third, located in the sequence position -605, is the first time that is determined. From the new proteins formed point of view, regarding the new allele 1, the four nucleotide changes in the sequence, imply only one change of amino acid in the protein.

This change is observed in exon 4, where there is a SNP-Y at position 219. This position is known as a common polymorphism site in the  $\alpha 3$  domain of HLA-G, and is related to a diversity of possible affinity between this domain and CD8 $\alpha\alpha$  and LILRBs. Regarding the protein formed by the new allele 2, the third nucleotide change, produces a non-silent substitution at position 178. This change results in a productive amino acid substitution from methionine (as in the rest of alleles heretofore described) to threonine, with a corresponding change in polarity. This polymorphic change, has not been described previously in any allele HLA-G.

## CONCLUSIONS.

- 1- New extended haplotypes HLA-G/HLA-clase I/HLA-clase II have been described for the first time in Amerindians.
- 2- It has been observed that, contrary to expectations, the frequencies of HLA-G alleles in Amerindians are more similar to European ones than to Asian ones.
- 3- High frequencies of HLA-G \* 01: 05N null allele, are recorded in populations of the Middle East, while they are very low in Amerindians. The environmental factors that have been able to maintain this great difference and their functional and evolutionary implications are discussed.
- 4- A new HLA-G allele has been found in an Ecuadorian population of the Altiplano: HLA-G \* 01: 08: 02. There are several changes in the molecule: in the  $\alpha 3$  domain (residue 219, Arg  $\rightarrow$  Trp), it could change the interaction of the new molecule with CD8 and the LILRBs receptors, that is, of cells presenting with B lymphocytes, T, NK, macrophages, Monocytes and dendritic cells.
- 5- A new HLA-G allele has been found in a Caribbean Sea population (Dominican Republic), HLA-G \* 01: 20, with a significant change in the  $\alpha 2$  domain (residue 178, Met  $\rightarrow$  Thr) at one corner of the loops of the HLA-G leaflet, which may affect the interaction with the clonotypic lymphocyte receptor.
- 6- HLA-G\*01:05N null alleles, have been observed in low frequency in Amerindian populations of Central America, North and South of South America.
- 7- Twelve new HLA-A-B-DRB1-G haplotypes have been described in Amerindians. In addition, four new haplotypes, without HLA-G, have been found in the Amerindian immigrant population in Madrid.
- 8- New molecules and new haplotypes that have been described here, may have relevance in the physiology and pathology associated with HLA and HLA-G in Amerindians.



# RESUMEN



## **INTRODUCCIÓN.**

El complejo mayor de histocompatibilidad (MHC, del inglés *Major Histocompatibility Complex*), constituye un conjunto de genes, en su mayoría altamente polimórficos, cuyos productos se expresan en la superficie de gran variedad de células y que son responsables de la respuesta inmune “adaptativa”. Se considera que estos genes están presentes en todos los vertebrados.

Estas moléculas son glicoproteínas de membrana, que actúan como presentadoras de antígeno a los linfocitos T. Se han definido clásicamente como antígenos de trasplante, ya que median las reacciones de rechazo que se producen contra los injertos de órganos. El sistema principal de histocompatibilidad del hombre se denomina sistema HLA (del inglés, Human Leukocyte Antigen).

Se localiza físicamente en el brazo corto del cromosoma 6 humano, ocupando una región de aproximadamente 4 millones de pares de bases (4Mb). La función principal de las moléculas HLA es facilitar el despliegue de fragmentos moleculares únicos en la superficie de las células, en un orden que permite su identificación mediante efectores inmunes, como los Linfocitos T. las moléculas HLA de clase I y clase II son glicoproteínas de membrana responsables de los pasos de reconocimiento y unión, mientras que otros genes de esta región, contribuyen en el procesamiento y presentación antigénico de otra forma.

Un gen polimórfico es aquel que tiene una alta frecuencia de variantes genéticas, y los genes HLA son los más polimórficos de la especie humana. Esto se debe a su principal papel en la respuesta inmune.

Los genes HLA de clase I codifican para proteínas formadas por una cadena pesada (cadena  $\alpha$ ), unida no covalentemente con la  $\beta$ 2-microglobulina. Los genes HLA de clase II, también son heterodímeros formados por una cadena  $\alpha$ , unida no covalentemente con una cadena  $\beta$ . Ambas proteínas, HLA clase I y clase II, presentan una estructura similar que consiste en una región extracelular, una región transmembrana y una cola citoplasmática. Las moléculas de clase I tienen tres dominios ( $\alpha$ 1,  $\alpha$ 2 y  $\alpha$ 3) de los cuales, los dominios  $\alpha$ 1 y  $\alpha$ 2, en la parte distal, forman la valva de unión al péptido. El dominio  $\alpha$ 3 junto con la  $\beta$ 2-m forman la región proximal. Las cadenas  $\alpha$  y  $\beta$  de las moléculas de clase II, tienen dos dominios cada una, la región distal  $\alpha$ 1 y  $\beta$ 1 forman la valva de unión al péptido, mientras que  $\alpha$ 2 y  $\beta$ 2-m forman la región proximal.

HLA-G es una molécula de clase I, no clásica cuya estructura es similar a las proteínas HLA de clase I. El gen HLA-G tiene un bajo polimorfismo, (53 alelos y 18 proteínas descritas hasta la fecha), una baja expresión tisular y la capacidad de formar dímeros. Por otro lado, mediante un proceso de *splicing* alternativo, se sintetizan siete isoformas distintas (HLA-G1 a HLA-G7), cuatro unidas a membrana (G1-G4), y tres solubles (G5-G7). Todas estas características contribuyen al aumento del interés científico en esta molécula, además de su importancia en la inducción de tolerancia inmunológica. En la molécula de HLA-G, el péptido se encuentra localizado más profundo dentro de la valva.

La expresión de HLA-G está restringida a tejidos específicos como tejido fetal, trofoblastos, timo, córnea y precursores eritrocíticos y endoteliales. Sin embargo, puede ser inducido en determinadas situaciones como trasplantes, enfermedades inflamatorias, células tumorales, infecciones virales etc.

El poblamiento de América y la procedencia de la población nativa americana, ha sido el objeto de numerosos estudios. Existen numerosas teorías acerca del poblamiento de América. Una de las más significativas es la teoría de Greenberg. Él postuló la teoría de la triple oleada migratoria: amerindios (entre 30.000- 15.000 años A.C), Na-Dene (atabascos, navajos y apaches, entre 15.000-10.000 A.C) y Esquimo-Aleutianos (entre 10.000-8.000 A.C). Por otro lado, también se ha sugerido la teoría transpacífica, desde Asia o Polinesia hasta América. Por último, la teoría transatlántica que sugiere la entrada desde Europa antes del descubrimiento de América por Colón, ya que, se han encontrado restos arqueológicos en Nuevo Méjico de herramientas usadas hace 20.000 años en España.

La mayoría de los inmigrantes amerindios de Madrid provenientes de países andinos (Ecuador, Colombia, Perú y Bolivia, principalmente), han llegado en los últimos 10 años. Este grupo poblacional presenta un perfil HLA con alelos casi específicos que los convierte en distintos al resto de poblaciones mundiales, y no han sido determinados sus alelos HLA-G anteriormente. Inmigrantes voluntarios sanos y no emparentados han participado en este estudio.

### **OBJETIVOS.**

- 1- Determinar alelos de HLA-G en una población de amerindios inmigrantes en Madrid.
- 2- Calcular los haplotipos extendidos de HLA-G/HLA-A/HLA-DRB1 (clase I y clase II) en la misma población.
- 3- Comparar las frecuencias de alelos HLA-G amerindios con los del resto de poblaciones mundiales.
- 4- Buscar posibles alelos nuevos del locus HLA-G en amerindios.
- 5- Extraer conclusiones funcionales y evolutivas sobre HLA-G de los resultados (principalmente del alelo nulo HLA-G\*01:05N), en las diferentes poblaciones del mundo.

### **RESULTADOS.**

Para un total de 154 individuos han sido determinados sus alelos HLA-G. Los individuos del estudio son amerindios inmigrantes en Madrid, voluntarios, sanos y no emparentados.

De los 53 alelos descritos hasta la fecha, solo 18 están representados en esta muestra. Los grupos más comunes son: HLA-G\*01:01 (71.8%), HLA-G\*01:04 (18.5%), HLA-G\*01:03 (5.8%) and HLA-G\*01:06 (2.9%). El haplotipo extendido más frecuente es HLA-A\*24:02-B\*35:01-DRB1\*14:02-G\*01:04, con una frecuencia de casi el 3%. Este haplotipo, sin el locus HLA-G, se ha encontrado también en poblaciones amerindias como aymaras (2.8%), quechuas (1.4%) y jaidukamas (7.7%).

Hay que mencionar que, en el resto de poblaciones mundiales, los haplotipos extendidos estudiados, no portan el locus HLA-G, por lo que cuando se hacen comparaciones con otras poblaciones, no tienen en cuenta este locus.

El segundo y tercer haplotipo extendido más frecuente, HLA-A\*02:01-B\*35:04-DRB1\*08:02-G\*01:01 (1.7%) y HLA-A\*02:01-B\*40:02-DRB1\*04:07-G\*01:01 (1.2%), también se presentan en otras poblaciones amerindias. Sin embargo, se puede apreciar cierta influencia europea, ya que, en el cuarto haplotipo extendido más frecuente, HLA-A\*03:01-B\*07:02-DRB1\*15:01-G\*01:01 (1,2%) con un posible origen europeo, se ha observado con frecuencias elevadas en poblaciones del norte de España e Irlanda.

Se han encontrado tres alelos nulos HLA-G\*01:05N, en tres individuos heterocigotos. La presencia de este alelo se ha observado con baja frecuencia (1-1.3%,). Esto también ocurre en poblaciones europeas caucásicas, orientales, pigmeos africanos y brasileños. En el caso de los grupos amerindios Mayas y Uros, la presencia de este alelo es nula. Tres haplotipos extendidos portan este alelo nulo en amerindios inmigrantes de Madrid, HLA-A\*33:03-B\*57:03-DRB1\*13:02-G\*01:05N, HLA-A\*31:01-B\*15:05-DRB1\*04:03-G\*01:05N y HLA-A\*29:02-B\*15:05-DRB1\*04:03-G\*01:05N, todos ellos con una baja frecuencia (0,6%) y se han observado en otras poblaciones. Solo se han encontrado diferencias significativas ( $p < 0.05$ ) entre la población amerindia inmigrante en Madrid y las poblaciones del Este Medio (Norte de India, iraníes e iraquíes) y negros de África, excepto pigmeos (africanos de Ghana y Shona y afroamericanos) ya que estos últimos presentan las frecuencias más elevadas para este alelo. Además, el alelo HLA-G\*01:05N presenta desequilibrio de ligamiento con el alelo HLA-A\*30:01 también en la población de estudio.

Durante el proceso de determinación del alelo HLA-G, se han encontrado dos alelos nuevos en individuos procedente de Ecuador (alelo nuevo 1) y República Dominicana (alelo nuevo2). En el caso del alelo nuevo 1, se han determinado los exones 2, 3 y 4, en los cuales se han encontrado cuatro cambios de nucleótidos, creando así una secuencia nueva. En relación al alelo nuevo 2, se han determinado los exones 2 y 3. Las variaciones nucleotídicas solo se han encontrado en el exón 3.

Respecto a las nuevas proteínas formadas, para el alelo nuevo 1, de los cuatro cambios producidos en la secuencia de ADN, solo uno provoca cambio de aa en la secuencia aminoacídica (residuo 219). Esto ocurre en el dominio  $\alpha 3$  de HLA-G, lugar de unión de esta molécula con los receptores LILRBs y CD8 $\alpha\alpha$ , por lo que podría afectar en dichas interacciones. En el alelo nuevo 2, solo el tercer cambio en la secuencia de ADN provoca un cambio de aa en la secuencia aminoacídica (residuo 178) en el dominio  $\alpha 2$ , lugar de interacción con el péptido y el TCR, por lo que también podría afectar a dicha interacción.

### **CONCLUSIONES.**

- 1- Se han descrito, por vez primera, nuevos haplotipos extendidos HLA-G/HLA-clase I/HLA-clase II en amerindios.
- 2- Se ha observado que, en contra de lo esperado, las frecuencias de alelos HLA-G en amerindios se asemejan más a las europeas que a las asiáticas.
- 3- Se registran unas frecuencias altas, de alelo nulo HLA-G\*01:05N, en poblaciones de Oriente Medio, mientras que son muy bajas en amerindios. Se discuten los factores ambientales que han podido mantener esa gran diferencia y sus implicaciones funcionales y evolutivas.
- 4- Se ha encontrado un alelo nuevo HLA-G en una población ecuatoriana del Altiplano: HLA-G\*01:08:02. Existen varios cambios en la molécula: en el dominio  $\alpha 3$  (residuo 219, Arg→Trp), podría cambiar la interacción de la molécula nueva con CD8 y los receptores LILRBs, es decir, de células presentadoras con linfocitos B, T, NK, macrófagos, monocitos y células dendríticas.
- 5- Se ha encontrado un alelo nuevo HLA-G en una población del Mar Caribe (República Dominicana), HLA-G\*01:20, con un cambio importante en el dominio  $\alpha 2$  (residuo 178, Met→Thr) en una esquina de los bucles de la valva de HLA-G que puede afectar a la interacción con el receptor clonotípico de los linfocitos.
- 6- Se han observado alelos nulos HLA-G\*01:05N en baja frecuencia en amerindios de Centroamérica, Norte y Sur de Sudamérica.
- 7- Se han descrito doce nuevos haplotipos HLA-A-B-DRB1-G en amerindios. Asimismo, se han encontrado cuatro nuevos haplotipos extendidos, sin HLA-G, en la población amerindia inmigrante en Madrid.
- 8- Las nuevas moléculas y los nuevos haplotipos que se han descrito, pueden tener una relevancia en la fisiología y patología asociada a HLA y a HLA-G en amerindios

# INTRODUCCIÓN





## **I. GENERALIDADES DEL SISTEMA HLA.**

### **1. DEFINICIÓN.**

El complejo mayor de histocompatibilidad (MHC, del inglés *Major Histocompatibility Complex*), o complejo principal de histocompatibilidad, constituye un conjunto de genes, en su mayoría altamente polimórficos, cuyos productos se expresan en la superficie de gran variedad de células y que son responsables de la respuesta inmune “adaptativa”. Se considera que estos genes están presentes en todos los vertebrados (Gomez-Casado et al., 2006) y fueron descubiertos en 1948 por Elwood Briles en el pollo (Briles & McGibbon, 1948)

Estas moléculas son glicoproteínas de membrana muy polimórficas que actúan como presentadoras de antígeno a los linfocitos T. Se han definido clásicamente como antígenos de trasplante, ya que median las reacciones de rechazo que se producen contra los injertos de órganos. El sistema principal de histocompatibilidad del hombre se denomina sistema HLA (del inglés, Human Leukocyte Antigen). Se localiza físicamente en el brazo corto del cromosoma 6 humano, ocupando una región de aproximadamente 4 millones de pares de bases (4Mb) (Arnaiz-Villena et al., 1995).

El sistema HLA es dialélico, codominante y presenta ciertas características. Como se ha referido con anterioridad, es un sistema altamente polimórfico, existen multitud de alelos descritos para cada locus, cada uno de ellos con distinta afinidad y unión por el antígeno proteico. De esta multitud de alelos, cada individuo tiene dos heredados, uno vía paterna y otra vía materna. Por otro lado, hay que destacar su carácter poligénico, es decir, varios genes organizados en tres regiones. Por último, este sistema presenta desequilibrio de ligamiento, es decir, diferentes alelos de distintos genes se encuentran en el mismo cromosoma con una frecuencia mayor a la teóricamente esperada en una combinación al azar.

Todas estas propiedades hacen que el MHC sea uno de los sistemas genéticos más complejos y a la vez más fascinantes descritos en la naturaleza.

## **2. HISTORIA.**

El intento de trasplantar fluidos, tejidos u órganos de unos seres humanos a otros, tiene una larga prehistoria difícil de datar. El alotrasplante (trasplante entre individuos de la misma especie) no empezó a tener una base científica hasta principios del siglo XX, cuando en 1901 Karl Landsteiner descubrió los diferentes grupos sanguíneos del sistema ABO. Así, se evidenció que el alotrasplante de sangre solo daba un resultado positivo cuando ambas sangres eran de idéntico grupo, aunque no fuese del todo exacto ya que existía un donante universal (grupo O) y un receptor universal (grupo AB).

Más tarde se comprobó que existía otro sistema de compatibilidad que afectaba a los demás tejidos orgánicos. El descubrimiento del sistema HLA ocurrió estudiando el rechazo de trasplantes. En 1902, Alexis Carrel puso a punto la técnica de sutura de vasos sanguíneos pero los trasplantes seguían siendo infructuosos (Gracia et al., 1996).

Fue en 1948 cuando fueron identificados antígenos eritrocitarios en el pollo, comprobándose que estaban codificados en 2 loci independientes (A y B) (Briles et al., 1950; Briles & McGibbon, 1948). En ese mismo año, George Snell descubrió el primer componente del Complejo Mayor de Histocompatibilidad a través del papel que ejercía en los rechazos de trasplantes en ratones (SNELL, 1948).

Años después, en 1958, Jean Dausset descubrió en la superficie de los glóbulos blancos el primer antígeno en humanos. Dausset observó que un paciente que recibía sucesivas transfusiones de sangre compatible con el sistema ABO, seguía sufriendo una inesperada reacción inmunológica. En este caso, detectó que los anticuerpos estaban atacando a los glóbulos blancos del donante. Él llamó a este antígeno MAC (correspondiente al HLA-A2) y fue el primero aislado en el sistema HLA (Dausset, 1958).

Al mismo tiempo Rose Payne y Jon Van Rood observaron anticuerpos que fueron detectados en el suero de mujeres multíparas que habían desarrollado inmunidad contra los antígenos leucocitarios paternos presentes en el feto (Payne & Rolfs, 1958; VAN ROOD et al., 1958).

En 1962, usando este suero Van Rood definió dos sistemas alélicos que él denominó grupo 4. Posteriormente esto fue desarrollado por Payne, quien definió un sistema aparentemente independiente de antígenos que él llamó LA (L de leucocitos y A del primer locus) (Bodmer, 1987).

Más tarde, se hizo evidente que estos antígenos pertenecían a dos series, cada una de las cuales se comportaban como si estuvieran controlados por un conjunto de alelos en

dos loci estrechamente vinculados, ahora conocidos como los loci HLA-A y HLA-B. Seguidamente un tercer locus fue definido, HLA-C (Bodmer, 1987).

Tras nuevos avances donde se determinó el locus HLA-DW, se estableció que este estaba en paralelo a un conjunto de determinantes serológicos que se podían identificar predominantemente en los linfocitos B de sangre periférica, el locus HLA-DR. Así, estudios celulares y serológicos identificaron adicionalmente tipos similares llamados HLA-DQ y DP (Bodmer, 1987).

El papel principal de los genes HLA en la respuesta inmune frente a antígenos fue postulada en 1970, cuando se demostró que los linfocitos T antígeno-específicos no reconocían antígenos libres o en forma soluble, sino que reconocían porciones de proteínas antigénicas unidas no covalentemente a las moléculas HLA (Gomez-Casado et al., 2006).

En definitiva, esto hizo deducir la importancia capital de estos antígenos en la defensa del organismo contra toda agresión externa o interna y, en consecuencia, responsable de los fenómenos de rechazo en un trasplante de órganos o de una transfusión sanguínea, lo que obligó al establecimiento de nuevas estrategias.

### **3. LOCALIZACIÓN GENÓMICA Y CLASIFICACIÓN.**

En 1999 fue publicado el primer mapa completo con las secuencias genómicas del complejo mayor de histocompatibilidad. El principal objetivo era producir las características genéticas y genómicas del mapa incorporando genes conocidos. Desde entonces, la plantilla de la secuencia genómica de MHC ha sido usada en numerosas ocasiones para investigar SNPs (en inglés, *Single Nucleotide Polymorphism*) y variaciones haplotípicas, expresión de genes, diversidad en las secuencias entre individuos de la misma y de diferente especie y la evolución de la organización estructural del MHC (Shiina et al., 2009).

La región MHC que abarca 4Mb, ocupa un 0.13% del genoma humano ( $3 \times 10^9$  bp), pero contiene aproximadamente un 0.5% de los 32.000 genes codificantes conocidos. Esta región es una de las más densas y mejor estudiadas dentro del genoma humano, y las secuencias aun por definir representan un porcentaje bajo de la región. (Shiina et al., 2009).

### 3.1 Localización.

Como se ha mencionado anteriormente, el sistema HLA se localiza físicamente en el brazo corto del cromosoma 6 humano, en la parte distal de la banda 6p21.3 y ocupa una longitud de 4 centimorgan (CM), aproximadamente 4 millones de pares de bases y contiene al menos 300 genes, pseudogenes y fragmentos génicos (Petersdorf, 2013).

Los genes identificados dentro de este complejo, dependiendo de la funcionalidad y estructura de sus productos, se dividieron en tres regiones principales (más tarde ampliadas): genes de clase I, en la zona telomérica del complejo y abarca un segmento cromosómico de unas 1600 kilobases (Kb). Genes de clase II, situados en la región centromérica y comprenden unas 900 Kb, y por último, genes de clase III, fuertemente ligados, abarcan un segmento de ADN de unas 100 Kb e incluyen las proteínas del complemento y el factor de necrosis tumoral (TNF) entre otros (Arnaiz-Villena et al., 1995). Debido a la identificación y caracterización constante de nuevos genes no se excluye una revisión futura de esta clasificación.

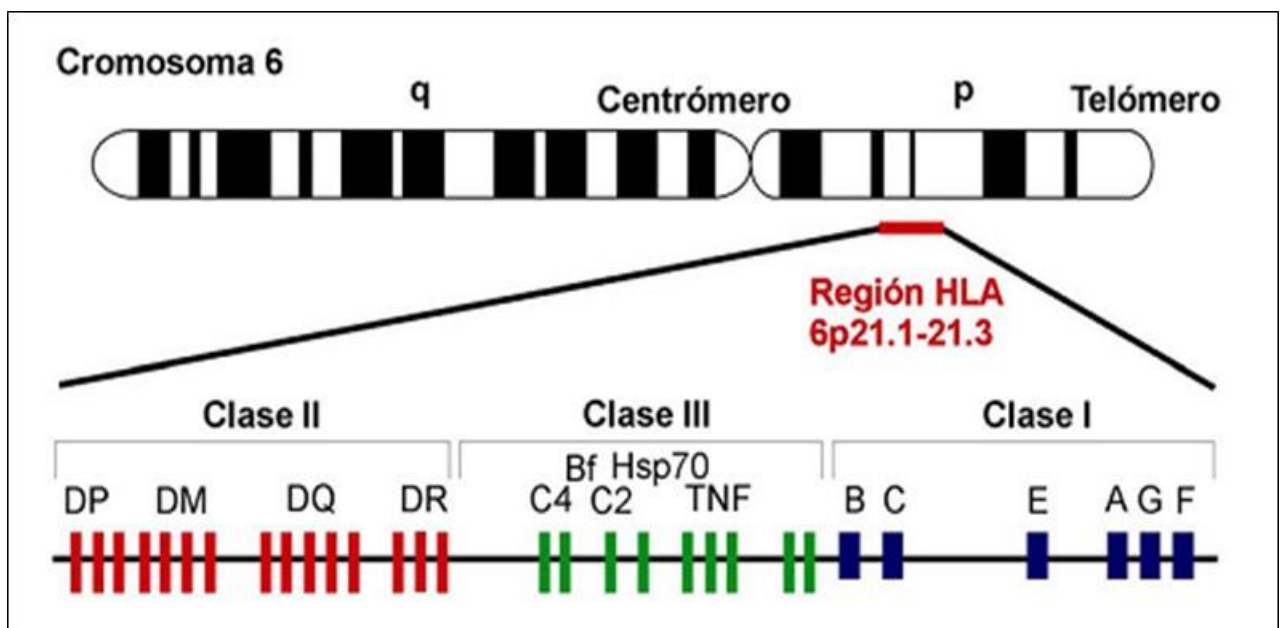


Figura I1: Mapa genético de la región HLA.

### 3.2 Clasificación.

El súper locus HLA es dividido tradicionalmente en 5 regiones que son: región clase I extendida, región clase I, región clase II, región clase II extendida y región clase III.

### **3.2.1 Región de clase I extendida:**

Es el segmento más telomérico. A esta región pertenecen genes que codifican para receptores olfatorios, histonas, ARNt (ARN transferente), proteínas dedos de zinc (en inglés *zinc-finger*), genes MOG, entre otros, y el gen de la hemocromatosis (HFE), estructuralmente parecido a los genes HLA de clase I, codifica para una proteína que parece estar implicada en el metabolismo del hierro. (Horton et al., 2004; Shiina et al., 2009).

### **3.2.2 Región de clase I:**

Esta región abarca desde HCP5P15 a MICB (Shiina et al., 2009). Comprende seis subgrupos de genes cuya nomenclatura está relacionada con el orden cronológico de su descripción, estructura, funcionalidad y patrón de expresión celular.

#### **- Genes HLA de clase I, clásicos o Ia:**

A este grupo pertenecen los genes HLA-A, -B y -C. Son los primeros descritos dentro del sistema HLA. Codifican para glicoproteínas de membrana que se expresan en prácticamente todas las células del organismo, si bien, su nivel de expresión varía desde un máximo en células pertenecientes al sistema inmune (linfocitos T, B y macrófagos) hasta un mínimo en células musculares, del sistema nervioso y fibroblastos. Son moléculas implicadas en la restricción del reconocimiento antigénico mediada por linfocitos T citotóxicos.

#### **- Genes HLA de clase I, no clásicos o Ib:**

A este grupo pertenecen los genes HLA-E, -F y -G. Codifican para proteínas estructuralmente similares a las de los genes clásicos. Se diferencian de los anteriores por su limitada expresión tisular, su bajo polimorfismo y por su función aún poco conocida.

Además, en la región de clase I se han descubierto una serie de secuencias de ADN, que se han clasificado en:

Pseudogenes: Nombrados como HLA-H, -J, -K y -L. Se sitúan en las proximidades del locus HLA-A, teloméricos o centroméricos a éste. Presentan deleciones que conllevan la aparición de codones de terminación prematuros. Se desconoce si tienen alguna función biológica (Geraghty et al., 1992; Robinson et al., 2015).

Genes truncados: Son los denominados HLA-N, -P, -S, -T, -U, -V, -W, -X. Presentan homología con los genes de clase I en zonas extensas. (Robinson et al., 2015).

#### **- Genes no HLA que se encuentran dentro de la región de clase I:**

Existe un grupo de genes dentro de la región de clase I que no tienen ninguna función conocida que los relacione con la presentación antigénica, algunos podrían ser genes reguladores. Aquí se encuentran genes de la familia P5, genes MOG y los más destacables los genes MIC, entre otros. (Shiina et al., 2009).

### **3.2.3 Región de clase II:**

Esta región contiene los genes de las cadenas alfa ( $\alpha$ ) y beta ( $\beta$ ) de la clase II. Situada en la zona centromérica del sistema HLA denominada región D. Se divide a su vez, en tres subregiones de centrómero a telómero: HLA-DP, -DQ y -DR. Los genes de clase II se definen con la letra D, seguida de la inicial de la subregión (P, Q o R) y de la letra A para los genes que codifican cadenas  $\alpha$ , o B para los genes que codifican cadenas  $\beta$ , dando lugar a las consiguientes subregiones, Además, se le añade un número para designar cuantos genes A o B están presentes en cada subregión (-DRA, -DRB1-B9, -DQA1, -DQB1, -DPA1, -DPB1). Existen aproximadamente 27 loci identificados en esta región, de los cuales 19 corresponderían al sistema HLA, 15 genes clásicos HLA y los 4 restantes, genes no clásicos HLA (-DOA, -DOB, -DMA y -DMB) que aparecen entre las subregiones -DP y -DQ. Dentro de los genes no pertenecientes al sistema HLA encontramos, entre otros, los genes TAP (TAP1 y TAP2) (en inglés, *transporter associated with antigen processing*) que codifican para proteínas transportadoras de péptidos y LMP (LMP2 y LMP7) (en inglés, *low molecular weight polypeptide*) (Casp et al., 2003), que intervienen en el procesamiento antigénico (Gomez-Casado et al., 2006; Robinson et al., 2015; Shiina et al., 2009).

### **3.2.4 Región de clase II extendida:**

Esta región, situada en la zona más centromérica, abarca solo 0,2 Mb desde el gen COL11A2 hasta KIFC1. Con 19 loci, de los cuales, 15 son genes codificantes, aunque para dos de ellos (WDR46 y PFDN6) era previamente desconocida su funcionalidad. (Shiina et al., 2009).

### **3.2.5 Región de clase III:**

Región localizada entre los genes de clase I y clase II, contiene 75 loci. En la región de clase III no todos los genes presentan función inmunológica, contiene genes de factores del complemento (C2, C4, CFB), y genes de citoquinas TNF (en inglés, *Tumor Necrosis Factor*), LTA, LTB (en inglés *Lymphotoxin-alfa* y *-beta*). Por otro lado, muchos de los productos de estos genes, tienen papeles fundamentales en procesos celulares como regulación transcripcional (CREBL1, PBX2...), mantenimiento (DOM3Z, NEU1...), biosíntesis, transporte de electrones y actividad hidrolasa

(DDAH2, ATP6V1G2...), interacción proteína-proteína para ambas interacciones, intracelular o intercelular, funciones de señalización y chaperonas (HSPA1A, TNXB...) (Shiina et al., 2009).

La presencia, en esta región, de genes sin relación funcional o evolutiva con los antígenos HLA, hace que muchos autores no la consideren como perteneciente al Sistema Principal de Histocompatibilidad (Gomez-Casado et al., 2006).

### **3.2.6 Repeticiones intercaladas:**

Aparte de los genes anteriormente comentados, el 49,5% de las secuencias genómicas del sistema HLA, están compuestas por elementos de repetición intercalados como por ejemplo, SINE (Alu, MIR), LINE (LINE1 y 2, L3/CR1), elementos LTR (ERVL, ERV clase I y clase II) y elementos ADN (*hAI-Charlie*, *TeMar-Tigger*) (Shiina et al., 2009).

## **4. ESTRUCTURA DEL SISTEMA HLA.**

### **4.1 Sistema HLA de clase I.**

#### **4.1.1 Estructura génica de los antígenos HLA de clase I, clásicos.**

Los genes que codifican para los antígenos HLA-A, -B y -C se caracterizan, como se ha comentado anteriormente, por su gran polimorfismo, por estar implicados en el rechazo de injertos y por expresarse de forma casi generalizada en las células del organismo.

Las moléculas clásicas presentan una cadena pesada  $\alpha$  codificada por un gen constituido por 8 exones, de manera que existe correspondencia exón/dominio estructural, excepto para los dominios citoplasmáticos que están codificados por dos o tres exones. El exón 1 codifica para la región promotora 5' UTR (del inglés, *UnTranslated Region*), y para un péptido líder de 24 aa, éste contiene un núcleo hidrofóbico característico de las proteínas que migran del Retículo Endoplasmático. Los exones 2 y 3 son los más polimórficos y codifican respectivamente para los dominios  $\alpha 1$  y  $\alpha 2$  de la molécula, mientras que el exón 4 que es menos polimórfico, da lugar al dominio  $\alpha 3$ . Los exones 5 y 6 codifican para el dominio transmembrana hidrofóbico. El dominio citoplasmático está codificado por los exones 6, 7 y parte del 8 para los genes HLA-A y -C, y por los exones 6 y 7 para HLA-B. En este último, el exón 8 ya codifica para la región 3'UTR (del inglés, *UnTranslated Region*). La diferencia en el tamaño del



dominio citoplasmático se debe a la situación que ocupa el codón de terminación, estando éste en el exón 8 para HLA-A y -C, y en el exón 7 para HLA-B (Gómez-Casado, 2002).

Por otro lado, las moléculas de clase I también están formadas por una cadena ligera llamada  $\beta 2$ -microglobulina ( $\beta 2$ -m) cuyo gen se encuentra fuera del sistema HLA, en el cromosoma 15 del ser humano. (Ploegh et al., 1981).

#### **4.1.2 Estructura proteica de los antígenos HLA de clase I, clásicos.**

Los antígenos HLA de clase I clásicos (HLA-A, -B y -C) son glicoproteínas de membrana compuestas por dos cadenas polipeptídicas, la cadena  $\alpha$  de unos 44 Kilodalton (Kd) y la  $\beta 2$ -m de 12 Kd. La cadena  $\alpha$  es el único miembro del heterodímero que atraviesa la membrana celular y cuyo extremo aminoterminal está orientado hacia el exterior de la célula. Ambas cadenas se unen no covalentemente en su porción extracelular (Gomez-Casado et al., 2006).

La cadena  $\alpha$  está formada por 338 aminoácidos (aa) en el caso de la molécula HLA-B y 341 aa en el caso de las moléculas HLA-A y -C. Su porción extracelular se divide en tres dominios globulares bien diferenciados de unos 90 aa cada uno, denominados: dominio  $\alpha 1$  (del aa 1 al 90, porción N-terminal de la cadena  $\alpha$ ), dominio  $\alpha 2$  (del aa 91 al 182) y dominio  $\alpha 3$  (del aa 183 al 247). Estos dominios están codificados por los exones 2, 3 y 4 respectivamente y pueden escindirse de la superficie celular bajo la acción proteolítica de la enzima papaína.

El dominio  $\alpha 3$  y la  $\beta 2$ -m presentan secuencias relativamente conservadas y de gran homología con los dominios constantes de las inmunoglobulinas. Por el contrario, los dominios  $\alpha 1$  y  $\alpha 2$ , fundamentalmente responsables del polimorfismo de los antígenos HLA de clase I, no presentan secuencias homólogas significativas con las regiones constantes o variables de las inmunoglobulinas (Corell et al., 1992). La porción transmembrana, con estructura de  $\alpha$ -hélice, de unos 27 aa, se continúa con un pequeño tallo citoplasmático de 30 aa aproximadamente, rico en tirosinas y serinas.

La cristalización de las moléculas es un paso previo a la determinación de estructuras tridimensionales mediante difracción de rayos X. En 1987, Bjorkman y colaboradores definieron la primera estructura tridimensional de una proteína HLA, concretamente la molécula HLA-A2, lo que originó que se multiplicasen los estudios sobre la relevancia funcional de los sitios de polimorfismo y su unión al receptor de células T (en inglés, *T cell Receptor*, *TCR*) (Bjorkman et al., 1987a). Todos los cristales

obtenidos presentaban una región distal formada por los dominios  $\alpha 1$  y  $\alpha 2$  de la cadena pesada, y una región proximal a la que pertenecen los dominios  $\alpha 3$  y  $\beta 2$ -m. (ver Figura I3.) (Corell et al., 1992; Gomez-Casado et al., 2006).

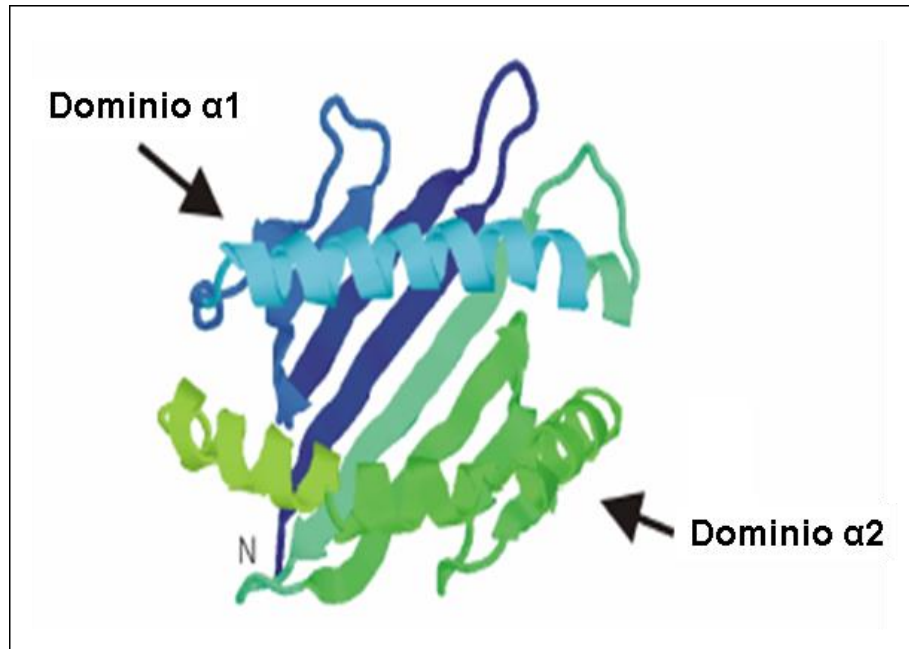


Figura I2: Representación tridimensional de la molécula HLA de clase I. Dominios moleculares ( $\alpha 1$  y  $\alpha 2$ ) de unión a antígeno. (Gomez-Casado et al., 2006)

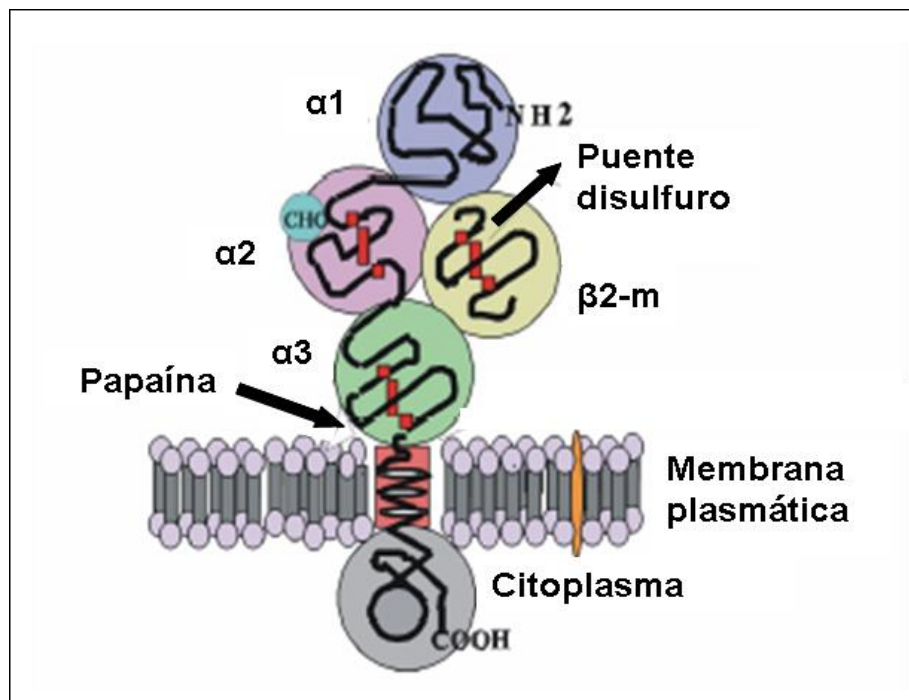


Figura I3: Estructura proteica de la molécula HLA de clase I. (Gomez-Casado et al., 2006).

Los dominios distales ( $\alpha 1$  y  $\alpha 2$ ) presentan una estructura compuesta por dos láminas  $\beta$ , formadas cada una por cuatro hebras antiparalelas, y una región de  $\alpha$ -hélice, zona más lejana a la membrana. Además, el dominio  $\alpha 2$  contiene una hélice corta adicional en posición carboxiterminal que une los dominios  $\alpha 2$  y  $\alpha 3$ . Las láminas  $\beta$  de ambos dominios forman la base de una valva flanqueada por las  $\alpha$ -hélices orientadas hacia el exterior. Esta valva, formada por los dominios  $\alpha 1$  y  $\alpha 2$  de las moléculas de clase I, constituye la estructura para la unión de los péptidos procesados (Moscoso, 2006).

El gran polimorfismo que muestran estas moléculas se encuentra concentrado en esta valva. Estos residuos polimórficos sirven para interactuar con los distintos péptidos y/o con el TCR. De los 20 residuos aminoacídicos de alta variabilidad identificados en las moléculas de HLA, 14 constituyen residuos de posible contacto con el péptido (posiciones 9, 24, 45, 66, 67, 70, 74, 77, 80, 95, 97, 114, 116 y 156), uno contacta con el TCR (posición 65) y cuatro más contactan con el péptido y con el TCR (posiciones 62, 69, 76 y 163) (Gomez-Casado et al., 2006).

Aparece un puente salino entre las posiciones 55 del dominio  $\alpha 1$  y 170 del dominio  $\alpha 2$ . Dos puentes disulfuro unen las cisteínas de las posiciones 101 y 164 de  $\alpha 2$  y las posiciones 203 y 259 de  $\alpha 3$ , además de varias interacciones de puentes salinos interdominios. Todas estas relaciones entre aa contribuyen a la estabilidad de la molécula de clase I en su parte distal. (Madden et al., 1992). En la posición 86 se localiza una asparagina N-glicosilada, pero esta ramificación no es esencial para el ensamblaje de la molécula ni para su expresión en superficie.

La región proximal a la membrana está formada por el dominio  $\alpha 3$  y la  $\beta 2$ -m y cada uno de ellos, al igual que los dominios constantes de las inmunoglobulinas, consiste en dos láminas  $\beta$  antiparalelas, una formada por cuatro hebras y la otra por tres. Las hebras de cada lámina se conectan entre sí por puentes disulfuro internos. La  $\beta 2$ -m se sitúa por la parte de abajo de la valva, de tal manera que existe un mínimo contacto entre estos dos dominios y el  $\alpha 3$  (Moscoso, 2006).

Si los dominios  $\alpha 1/\alpha 2$  interactúan con el péptido y con el TCR reconociendo residuos polimórficos, el dominio  $\alpha 3$  lo hace con el correceptor de la célula T (CD8), reconociendo determinantes monomórficos en  $\alpha 3$ . La interacción con CD8 requiere la presencia de aa ácidos en las moléculas de clase I, en posiciones 223, 227 y 229, que ocupan una zona localizada en un lazo del dominio  $\alpha 3$ , en una cavidad justo debajo de la valva (Connolly et al., 1990).

Existe una gran variabilidad entre las moléculas de clase I como consecuencia de cambios topológicos en los sitios de unión al péptido. Las características químicas y la exclusiva configuración de la estructura de cada valva (o *Peptide-Binding Groove*) explica cómo pueden unir gran variedad de péptidos de entre 8 y 10 aa generalmente endógenos (Janeway et al., 2003). En 1991, se identificaron una serie de depresiones a lo largo de la valva, conocidas como “*Peptide-binding Pockets*” o bolsillos (Saper et al., 1991). Los péptidos unidos a la valva pueden acomodar una o más de sus cadenas laterales aminoacídicas en estos bolsillos. Se han identificado seis bolsillos denominados con letras de la A a la F, localizados en las uniones de las láminas  $\beta$  con las  $\alpha$ -hélices (bolsillos B, C, D y E) o en los extremos de las dos  $\alpha$ -hélices (bolsillos A y F). Estos dos últimos contienen aa más conservados que los otros cuatro. De las 18 posiciones de alta variabilidad identificadas en las proteínas de clase I que interaccionan con el péptido y/o con el TCR, 16 están situadas en la valva (Barber & Parham, 1993; Gomez-Casado et al., 2006; Moscoso, 2006).

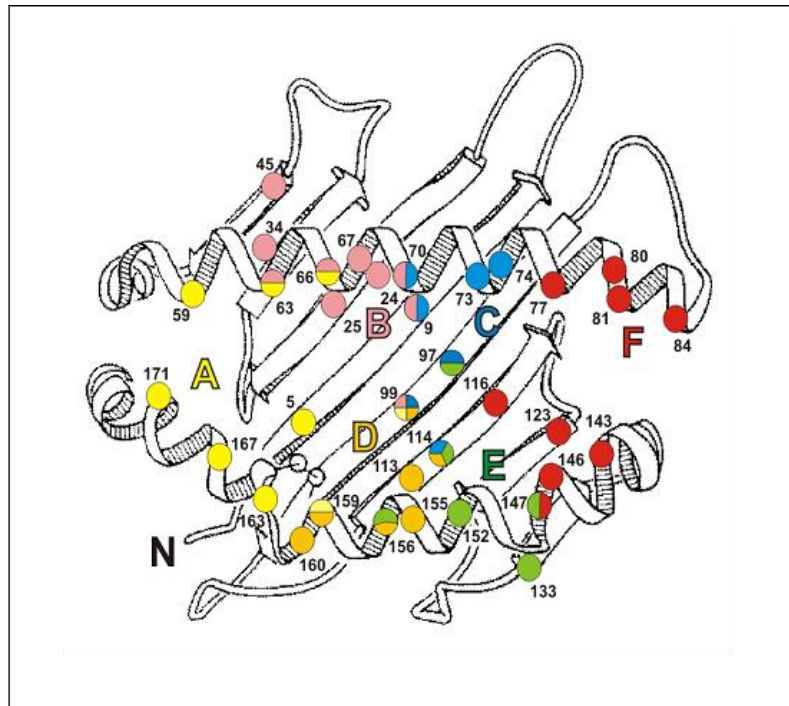


Figura I4: Disposición de los bolsillos de unión al péptido en la molécula HLA de clase I. El color de cada círculo (residuos variables) representa la contribución a cada bolsillo. (Gomez-Casado et al., 2006)

Las moléculas de clase I se consideran miembros de la superfamilia de las inmunoglobulinas por tener un origen evolutivo común y por similitudes estructurales, más evidentes en el dominio  $\alpha 3$ .

#### **4.1.3 Estructura génica de los antígenos HLA de clase I, no clásicos.**

No existen grandes diferencias en cuanto a la organización de exones e intrones de estos genes con respecto a los clásicos, aunque poseen una serie de características peculiares que los diferencian de estos últimos, anteriormente mencionadas.

HLA-E: Localizado genéticamente entre los loci HLA-C y HLA-A, fue descubierto en 1988. Tiene bajo polimorfismo. (Mizuno et al., 1988). La estructura genética es igual a las moléculas clásicas.

HLA-F: Fue caracterizada como gen de clase I no clásica en 1990 (Geraghty et al., 1990). Al igual que los otros genes no clásicos, HLA-F presenta una organización exón/intrón y tamaño similar a los clásicos. Posee dos características importantes: la primera de ellas se debe a una mutación puntual en 3' del intrón 6, que elimina por “splicing” (maduración del ARNm, ARN mensajero) el exón 7, resultando en un dominio citoplasmático más corto. La segunda es debida a un inusual extremo 3'UTR, exclusivo entre los genes de clase I, que se caracteriza por la inserción de una secuencia altamente conservada durante la evolución y semejante a la de una proteína ribosomal (Gomez-Casado et al., 2006).

HLA-G: Está codificada también por 8 exones, de los cuales, el exón 1 codifica para el péptido líder, el exón 2 y 3 codifican para los dominios  $\alpha 1$  y  $\alpha 2$  respectivamente, el exón 4 codifica para el dominio  $\alpha 3$ , exón 5 para la región transmembrana y por último el exón 6 que codifica para la cola citoplasmática. (NCBI, 2015a). Este corte es debido a un codón de terminación en el exón 6, por lo que el exón 7 está siempre ausente del ARNm maduro y el exón 8 no se traduce (Teixeira et al., 2013).

#### **4.1.4 Estructura proteica de los antígenos HLA de clase I, no clásicos.**

- HLA-E: Sus ARNm se han visto en todo tipo de tejidos, incluyendo tejidos fetales y placentarios (Gomez-Casado et al., 2006), además los transcritos de ARN se expresan en el citoplasma y superficie celular de citotrofoblastos en los últimos estadios del embarazo y sus niveles de expresión pueden verse aumentados con el interferón gamma (INF- $\gamma$ ) (Gomez-Prieto et al., 2010). En el análisis cristalográfico de HLA-E, se mostró como su estructura era similar a las moléculas clásicas (O'Callaghan et al., 1998). Es un

heterodímero formado por una cadena pesada  $\alpha$  y la  $\beta$ 2-m. La cadena pesada está anclada a la membrana. HLA-E une un conjunto restringido de péptidos derivados del péptido líder del resto de moléculas de clase I. La cadena  $\alpha$  está codificada por 8 exones. El exón 1 codifica para el péptido líder, el exón 2 y 3 para el dominio  $\alpha$ 1 y  $\alpha$ 2, ambos unen el péptido, el exón 4 codifica para el dominio  $\alpha$ 3, el exón 5 para la región transmembrana y por último los exones 6 y 7 que codifican para la cola citoplasmática (NCBI, 2015c).

- HLA-F: Es un heterodímero formado por una cadena pesada  $\alpha$  (anclada a la membrana) y la  $\beta$ 2-m. También presenta bajo polimorfismo (NCBI, 2015d).

Su expresión está restringida a células B, timo, bazo, amígdalas e hígado fetal. También se expresa en el citoplasma de trofoblastos (semanalmente en los tres primeros meses de embarazo) (Gomez-Prieto et al., 2010).

A diferencia de la mayoría de las otras cadenas pesadas  $\alpha$ , esta molécula se localiza en el retículo endoplasmático y en el aparato de Golgi, con una pequeña cantidad en la superficie de determinados tipos celulares (NCBI, 2015d).

HLA-F une un grupo restringido de péptidos para la presentación antigénica.(NCBI, 2015d). Posee residuos conservados de cisteína (posiciones 101 y 164) que son cruciales para conformar su estructura tridimensional y también el importante sitio de glicosilación, Asparagina 86. Por otro lado, conserva la mayoría de residuos esenciales a la hora de unir la molécula CD8. Su función es desconocida (Gomez-Prieto et al., 2010).

- HLA-G: Esta molécula de clase I no clásica es un heterodímero formado por la cadena pesada  $\alpha$  y la  $\beta$ 2-m. La cadena pesada  $\alpha$  está anclada a la membrana.

Esta molécula también posee un reducido polimorfismo, splicing alternativo en su transcrito primario, lo que da lugar a 7 isoformas proteicas, distribución limitada en tejidos y propiedades inmunosupresoras (Teixeira et al., 2013). Más adelante se explican en profundidad las características fundamentales de HLA-G tanto a nivel genético como molecular.



## **4.2 Sistema HLA de clase II.**

### **4.2.1 Estructura génica de los antígenos HLA de clase II.**

Los genes de clase II se localizan en la región más centromérica. Como ya se ha comentado, se divide a su vez en tres subregiones, de centrómero a telómero: -DP, -DQ y -DR.

La organización exón/intrón de los genes de clase II es similar a la organización de los genes de clase I, dado que los exones codifican por separado, para los dominios de la proteína (Marsh et al., 1999).

Por lo general, la cadena  $\alpha$  del gen de HLA de clase II, consta de 5 exones: el exón 1 codifica para la secuencia líder que contiene una región promotora 5', la secuencia señal y además codifica los primeros 5 aa de la proteína. Dos exones, el 2 y el 3, codifican para los dominios  $\alpha 1$  y  $\alpha 2$ , mientras el péptido de conexión y los dominios transmembrana y citoplasmático son codificados por el exón 4. La mayor parte de la región 3'UTR es codificada por el exón 5. La estructura del gen de la cadena  $\beta$  es muy similar, sin embargo, está constituida por 6 exones. El péptido líder está codificado por el exón 1, el exón 2 y 3 para los dominios de proteína  $\beta 1$  y  $\beta 2$  respectivamente, el exón 4 para el péptido conector, el dominio transmembrana, parte del dominio citoplasmático y el resto de este dominio está originado por, al menos, dos exones más, si bien esto último varía en las diferentes subregiones. (Gómez-Casado, 2002; Muro et al., 2015).

Subregión -DR: Dependiendo del haplotipo (genes o alelos que se heredan conjuntamente), aparece un solo gen DRA que codifica para las cadenas  $\alpha$  (HLA-DRA), cinco genes para las cadenas  $\beta$  (HLA-DRB1-5), y 4 pseudogenes (HLA-DRB6-9) (Muro et al., 2015). No todos los genes HLA-DRB están presentes en todos los haplotipos, de modo que algunas combinaciones son específicas, pueden ocurrir varios reordenamientos dentro de este locus. El gen DRB1, es el más polimórfico. (Gómez-Casado, 2002).

Subregión -DQ: Consta de dos genes para la cadena  $\alpha$ ; DQA1, -DQA2 y tres para las cadenas  $\beta$ , -DQB1, -DQB2 y -DQB3. Los productos de los genes HLA-DQA1 y -DQB1 se asocian para formar los antígenos HLA-DQ (Muro et al., 2015). Los demás son pseudogenes.

Los dos genes codificantes se orientan del mismo modo que los genes -DP. HLA-DQA1 presenta la estructura típica de los genes A de clase II. Sin embargo, -DQB1

tiene un dominio citoplasmático más corto al eliminarse por maduración de su ARNm, el exón 5 (Radley et al., 1994).

Subregión -DP: Ocupa unas 100 Kb y contiene los siguientes genes para la cadena  $\alpha$ ; -DPA1, -DPA2 y los genes para las cadenas  $\beta$ ; -DPB1 y -DPB2. Los genes -DPA1 y -DPB1, que se orientan cabeza con cabeza con sus promotores separados por 2 Kb, codifican para un heterodímero típico de clase II. -DPA2 y -DPB2, no se expresan, por lo que se consideran pseudogenes. Las causas de esta no expresión son desconocidas (Gómez-Casado, 2002).

Otros genes HLA de clase II: Entre las subregiones -DP y -DQ hay toda una serie de genes cuya estructura y función empieza a conocerse en la actualidad. Entre ellos están los genes -DOA, -DOB, -DMA y -DMB.

Los genes -DMA, -DMB y -DOA se sitúan más próximos a la subregión -DP y el gen -DOB lo hace cerca de -DQ (Horton et al., 2004).

El gen -DM (A y B) codifica para las cadenas  $\alpha$  y  $\beta$  de la proteína, no es polimórfico y se encuentra implicado en el procesamiento y la presentación antigénica. Cada uno de sus genes está estructurado en secuencias de intrones y exones relacionados con los dominios estructurales de la proteína codificada por ellos.

El gen -DO (A y B) también codifica para las cadenas  $\alpha$  y  $\beta$  formando un heterodímero. No es polimórfico. Su cadena  $\beta$  contiene 6 exones, al igual que el resto de genes de clase II. El exón 1 codifica para el péptido líder, los exones 2 y 3 codifican para los dos dominios extracelulares, el exón 4 para el dominio transmembrana y el exón 5 para la cola citoplasmática.(NCBI, 2015b).

#### **4.2.2 Estructura proteica de los antígenos HLA de clase II.**

Las moléculas de clase II (HLA-DP, -DQ y -DR) son también glicoproteínas de membrana de tipo I, forman heterodímeros transmembrana que constan de dos cadenas, una cadena  $\alpha$  de 33-35 (Kd) y otra  $\beta$  de 26-28 Kd asociadas no covalentemente (Gomez-Casado et al., 2006). Esta diferencia de tamaño se debe a la cantidad de carbohidratos de cada cadena. Cada una de ellas, consta de una región citoplásmica de 12 a 15 aa, una región transmembrana de 20-25 aa, y una región extracelular formada por dos dominios de unos 90-100 aa. Los dominios próximos a la membrana de cada cadena,  $\alpha 2$  y  $\beta 2$ , poseen una estructura del tipo de la superfamilia de las inmunoglobulinas, mientras que los dominios distales,  $\alpha 1$  y  $\beta 1$ , forman la estructura de unión al péptido.



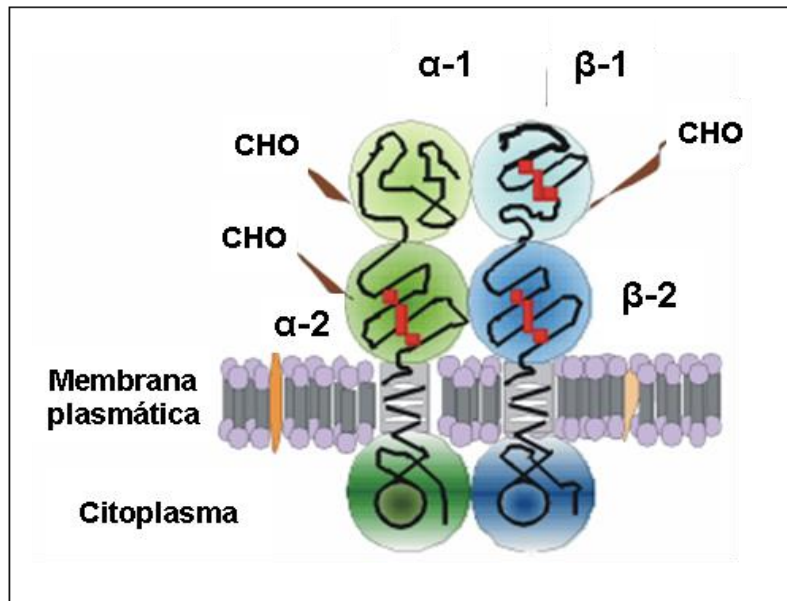


Figura I5: Estructura proteica de la molécula HLA de clase II. En rojo se marcan los puentes disulfuro (Gomez-Casado et al., 2006).

Las cadenas  $\alpha$  contienen dos puntos de N-glicosilación, uno en cada dominio extracelular y un puente disulfuro intradominio en  $\alpha 2$ , mientras que las cadenas  $\beta$  solo presentan un sitio de N-glicosilación en  $\beta 1$  y dos puentes disulfuro intradominio, uno en  $\beta 1$ , que genera un asa de 64 aa, y otro en  $\beta 2$ . (Corell et al., 1992).

Intracelularmente las proteínas de clase II se asocian con un tercer elemento proteico no polimórfico, de 31 Kd, denominado Cadena Invariante y codificado fuera del sistema HLA (Corell et al., 1992; Gomez-Casado et al., 2006). Este complejo formado por la Cadena Invariante y el heterodímero se forma durante la biosíntesis y transporte de las moléculas de clase II; sin embargo, se disocian antes de su expresión en la superficie celular. La Cadena Invariante podría jugar un papel importante en el transporte intracelular de las moléculas de clase II y en la unión del péptido (Kvist et al., 1982).

En 1993, se consiguió obtener la estructura tridimensional de la molécula HLA-DR1 gracias a las técnicas de difracción de rayos X. Pudieron observar que esta molécula era muy similar a las de clase I, de tal forma, los dominios  $\alpha 1$  y  $\alpha 2$  de HLA-DR1 se asemejaban al dominio  $\alpha 1$  y  $\beta 2$ -m de las moléculas de clase I y los dominios  $\beta 1$  y  $\beta 2$  de HLA-DR1 eran superponibles respectivamente a los dominios  $\alpha 2$  y  $\alpha 3$  (Brown et al., 1993).

El punto de interacción con CD4 se ubica en el dominio  $\beta 2$  de la molécula de clase II (Konig et al., 1992).

Los dominios distales de la proteína, ( $\alpha 1$  y  $\beta 1$ ), forman la valva de unión del péptido que consiste en una lámina  $\beta$  formada ocho hebras antiparalelas y flanqueada por dos regiones en  $\alpha$ -hélice. La región helicoidal del dominio  $\alpha 1$  presenta un extremo amino terminal en cadena extendida y su extremo carboxiterminal se pliega hacia la base de la valva. Esta peculiaridad hace que la zona de unión al péptido muestre los extremos abiertos. En esta valva abierta por los extremos se puede acomodar un péptido grande, de entre 12 y 24 aa. Este péptido puede permanecer en la valva en conformación extendida y sus extremos amino y carboxiterminales sobresalir fuera de la molécula (Rammensee et al., 1995; Rudensky et al., 1991).

El polimorfismo de estos genes se concentra en el exón 2, que es el responsable de codificar la estructura de la unión al péptido. Las regiones hipervariables, para las moléculas HLA-DRB1, se sitúan entre las posiciones aminoacídicas 25 a 40 y 65 a 80, y son las regiones diana de los métodos de determinación de alelos usados para definir los polimorfismos en el ADN mediante PCR (del inglés, *Polymerase Chain Reaction*) (Saiki et al., 1988). Las moléculas HLA-DQ y -DP son polimórficas en ambas cadenas  $\alpha$  y  $\beta$ . (Muro et al., 2015).

### Otros genes HLA de clase II

Las moléculas -DO forman un heterodímero  $\alpha/\beta$ . No son polimórficas, se encuentran en los lisosomas de los linfocitos B y suprimen la carga de péptidos de las moléculas de HLA de clase II por inhibición de HLA-DM (van Ham et al., 1997).

Las moléculas -DMA y -DMB también se organizan en un heterodímero  $\alpha/\beta$ , de membrana, con 263 aa aproximadamente, no aparecen en la superficie celular. La estructura proteica de este dímero sería ligeramente diferente a la del resto de los antígenos de clase II; los dominios proximales a la membrana pertenecen a la superfamilia de las inmunoglobulinas y los dos dominios distales contienen aa que pueden formar múltiples puentes disulfuro que darían lugar a una conformación muy rígida y probablemente cerrada. Estos genes poseen una función relacionada con la presentación antigénica mediante moléculas HLA de clase II (Gomez-Casado et al., 2006).

## **5. NOMENCLATURA.**

El desarrollo del sistema HLA ha sido mucho mayor desde que se iniciaron una serie de talleres internacional, llamados Workshops Internacionales de Histocompatibilidad, iniciados por Amos en 1964. Los trabajos de los talleres definen el desarrollo del sistema HLA (Bodmer, 1987). En el primer encuentro internacional se puso de relieve las dificultades de este tipo de trabajo de colaboración masiva. Los diferentes métodos experimentales, la inconsistencia en la ejecución de las mismas pruebas y una falta de homogeneidad de los sistemas de asignación de nombres, hacían necesaria una colaboración internacional.

El sistema de nomenclatura mas reciente para el sistema HLA, se desarrolló en el 2010 por el Comité para Factores del Sistema HLA de la OMS (Organización Mundial de la Salud) que se reunieron por primera vez en 1968, dejando cerrados los criterios para las sucesivas reuniones. Este comité se reúne regularmente para discutir cuestiones de nomenclatura y ha publicado 19 informes importantes que documentan, en primer lugar, los antígenos HLA y, más recientemente, los genes y alelos (Robinson et al., 2015).

Desde 1989, cuando un gran número de secuencias de alelos HLA se analizaron y se nombraron, el trabajo de mantenimiento de una base de datos de secuencias se hizo imprescindible (Robinson et al., 2015).

De tal manera, la nomenclatura actual se rige por las siguientes normas:

Cada nombre de un alelo HLA tiene un número único de cuatro, seis u ocho dígitos. La longitud de la designación del alelo depende de la secuencia de éste. Todos los alelos reciben, por lo menos, un nombre de cuatro cifras; el sexto y octavo dígito se asignan solamente cuando son necesarios.

Los primeros dos dígitos describen el tipo, que corresponde a menudo al antígeno serológico, el tercer y cuarto dígito se utilizan para enumerar los subtipos (alelos, propiamente dichos), números que son asignados de acuerdo con el orden en que se han determinado las secuencias de ADN. Los alelos cuyos números diferencian en los primeros cuatro dígitos deben diferenciarse en una o más sustituciones de nucleótidos que cambien la secuencia de aa de la proteína codificada. Los alelos que se diferencian solamente por sustituciones de nucleótidos sinónimas, (no cambia el aa para el que codifica) dentro de la región codificante se diferencian con el uso del quinto y sexto dígito. Alelos que solo se diferencian por cambios en la región no codificante (intrones)

o en las regiones 5' o 3' no traducidas que flanquean los exones y los intrones, son distinguidas por el uso del séptimo y octavo dígito.

Además del número único del alelo, hay sufijos adicionales que se pueden agregar a un alelo para indicar su expresión. Los alelos que se ha demostrado que no se expresan, se les añade el sufijo "N". Para indicar que un alelo tiene baja expresión en la superficie de la célula en comparación a los niveles normales se utiliza el sufijo "L". El sufijo "S" ayuda a denotar un alelo que especifica una proteína que es soluble, no está presente en la superficie de la célula. Una "C" añadida como sufijo indica que el producto del alelo está presente en el citoplasma, pero no en la superficie de la célula. El sufijo "A" indica una expresión aberrante, donde hay una cierta duda de si una proteína se expresa o no. La "Q" es agregada como sufijo cuando la expresión de un alelo es cuestionable, dado que la mutación considerada en el alelo se ha demostrado previamente que afecta a la expresión normal.(Robinson et al., 2015).

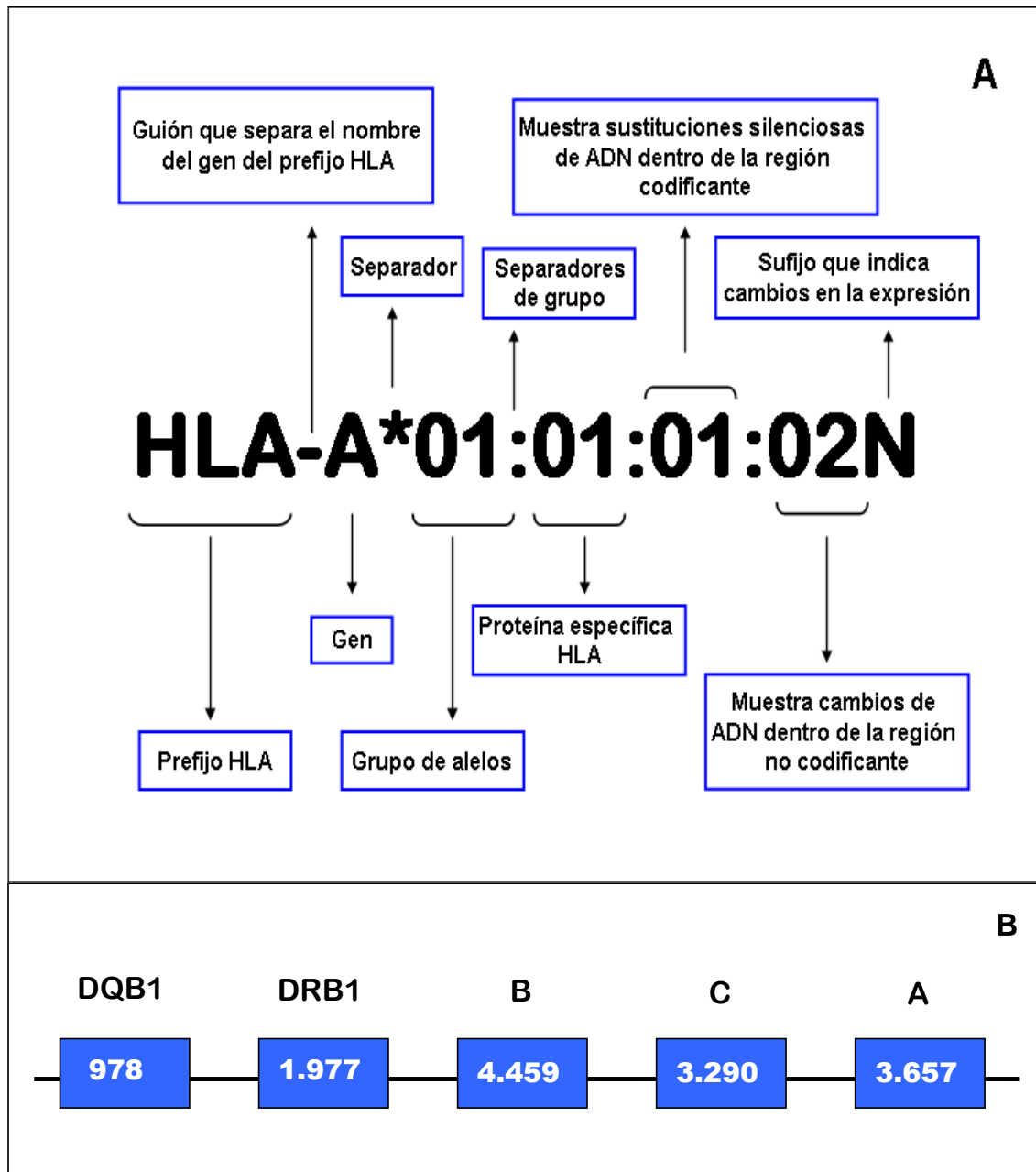


Figura I6: A) Esquema de la nomenclatura actual del sistema HLA. (Robinson et al., 2015). B) El gran polimorfismo observado hoy en la región HLA es fundamental para entender: a- su utilidad en el estudio de la genética de poblaciones; b- la dificultad de encontrar dos individuos idénticos para efectuar trasplantes; c- su utilidad en el estudio de enfermedades ligadas a HLA en las distintas poblaciones; d- la importancia del estudio farmacogenómico prospectivo (industria farmacéutica) y diagnóstico/pronóstico de los efectos de los fármacos ligados a HLA en diferentes poblaciones.

## 6. FUNCIÓN DE LAS MOLÉCULAS HLA: PRESENTACIÓN ANTIGÉNICA.

La función protectora de los linfocitos T depende de su capacidad para reconocer células que presentan agentes patógenos o que han internalizado agentes patógenos o sus productos. Los linfocitos T llevan a cabo dicha función reconociendo fragmentos

peptídicos de proteínas derivadas de agentes patógenos, en forma de complejos de péptidos y moléculas de HLA en la superficie de esas células. La generación de péptidos a partir de un antígeno intacto conlleva la modificación de la proteína nativa, y suele denominarse procesamiento antigénico, mientras que el acto de desplegar el péptido en la superficie celular por moléculas de HLA se denomina presentación antigénica (Janeway et al., 2003).

Las moléculas HLA de clase I y II corresponden a dos vías diferentes de procesamiento de los antígenos, y se asocian con dos sistemas diferentes de defensa inmunitaria.

### 6.1 Moléculas HLA de clase I.

Las moléculas de clase I actúan como receptores de péptidos endógenos, propios, normales, alterados (tumORAles) y virales, para hacerlos accesibles a la detección por las linfocitos T que llevan en su superficie celular el marcador CD8+ (mayoritariamente linfocitos citotóxicos) y destruyan la célula infectada (Gomez-Casado et al., 2006). En la siguiente figura se muestra el proceso de presentación antigénica.

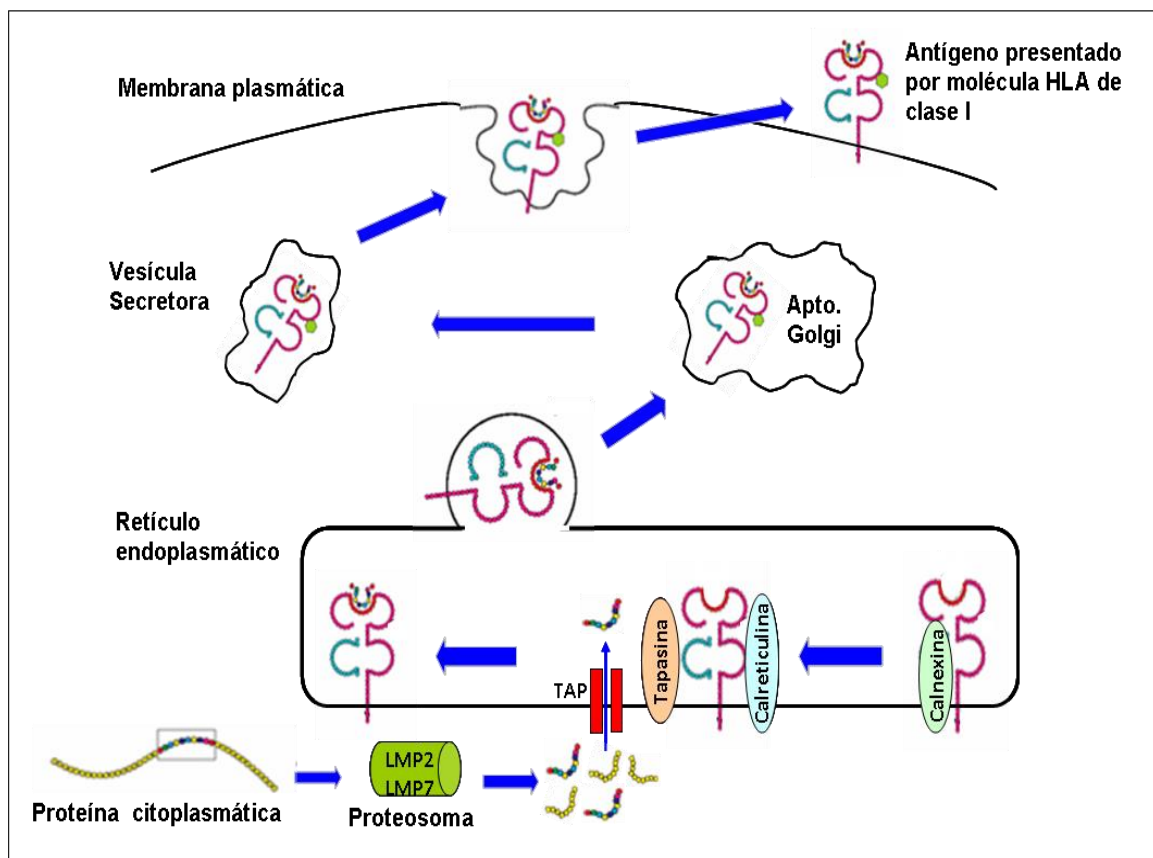


Figura I7: Vía de presentación antigénica endógena. Las moléculas de clase I actúan como receptores de péptidos endógenos para hacerlos accesibles a la las detección por linfocitos T

*CD8+ y destruyan la célula infectada (Gomez-Casado et al., 2006). Estas moléculas presentan selectividad alelo-específica (Falk et al., 1991). Todos los péptidos producidos en el citosol son translocadas al lumen del retículo endoplasmático, donde deben plegarse correctamente. Como se muestra en la figura, los péptidos presentados son degradados por proteasomas (Koopmann et al., 1997), formados por dos subunidades: LMP2 y LMP7 (Janeway et al., 2003). Posteriormente son traslocados a través de la membrana del retículo endoplasmático por los transportadores de péptidos TAP1 y TAP2. Dicho complejo muestra preferencia por los péptidos de 8 o 10 aa (Janeway et al., 2003). La cadena  $\alpha$  de la molécula HLA de clase I recién sintetizada, se une a **calnexina**, que retiene a la molécula en un estado parcialmente plegado en el retículo endoplasmático. Cuando la  $\beta$ 2-m se une a la cadena  $\alpha$ , ambas se disocian de la calnexina y se unen a un complejo proteico, formado por la **calreticulina** y un segundo componente que es la **tapasina**, permitiendo que la cadena  $\alpha$  y la  $\beta$ 2-m aguarden el transporte de un péptido apropiado desde el citosol. Finalmente, la unión de un péptido al heterodímero, cadena  $\alpha$ - $\beta$ 2-m parcialmente plegado, lo libera del complejo TAP-tapasina-calreticulina, permitiendo a la molécula HLA de clase I ser transportada hacia la membrana celular a través del cis y trans-Golgi. (Carreno et al., 1995; Janeway et al., 2003). La mayor parte de las interacciones con el péptido confieren estabilidad a la molécula y al péptido, y los hacen más accesibles al contacto con el TCR del linfocito T (Gomez-Casado et al., 2006).*

## **6.2 Moléculas HLA de clase II.**

Las moléculas de clase II unen un grupo heterogéneo de péptidos procedentes, en su mayor parte, del exterior celular y se encargan de exponerlos en la superficie de las células presentadoras de antígenos (linfocitos B, células dendríticas y macrófagos), ahí son reconocidos por el TCR de los linfocitos T CD4+ (generalmente cooperadores). La mayor parte de los péptidos presentados por las moléculas de clase II, provienen de la degradación endocítica de proteínas exógenas o bien, pero en menor medida, de proteínas endógenas que acceden a los endosomas. En la siguiente figura se muestra el proceso de presentación antigénica.

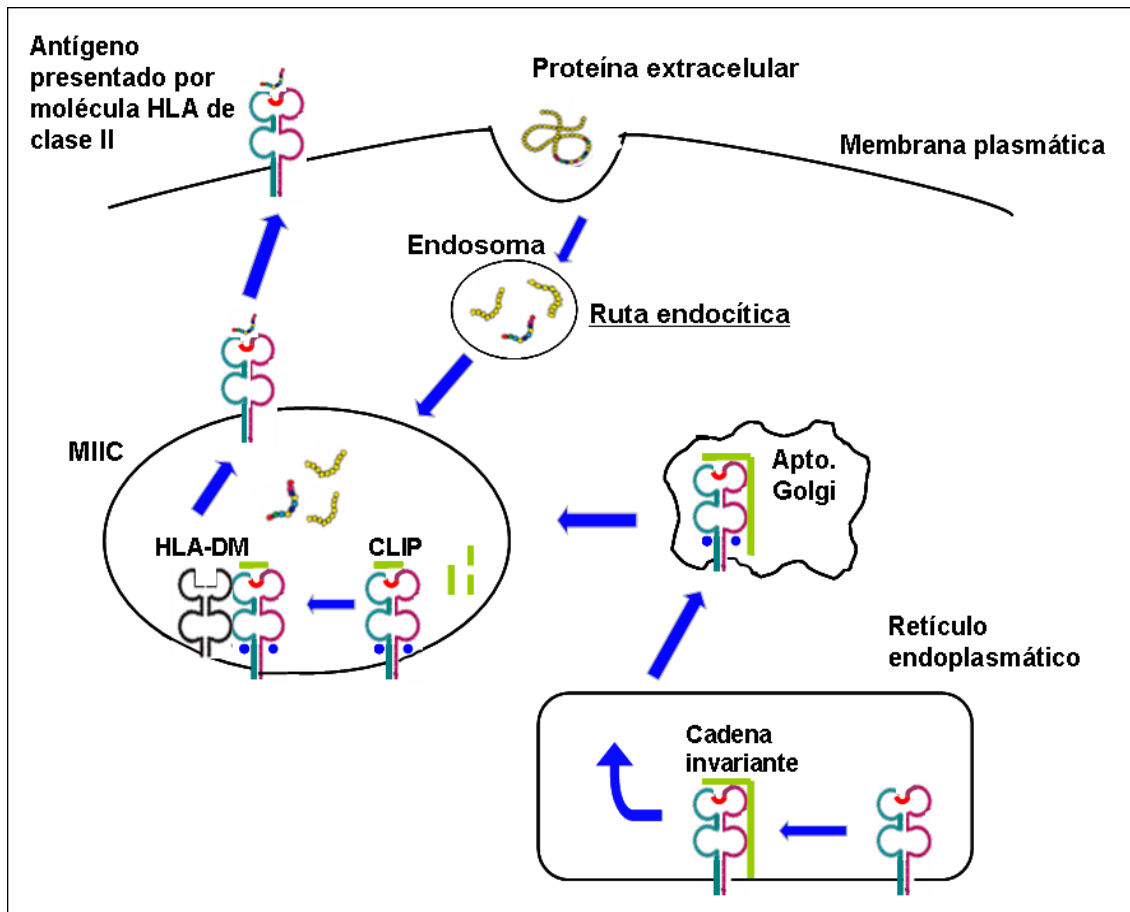


Figura I8: Vía de presentación antigénica exógena. Los antígenos interiorizados que penetran en la célula por endocitosis, quedan envueltos en el endosoma, que se acidifica de forma progresiva y degrada las proteínas contenidas en ellos. Las cadenas  $\alpha$  y  $\beta$  de los antígenos HLA de clase II son sintetizados en el retículo endoplasmático, el heterodímero naciente se asocia a la **Cadena Invariante** (proteína de membrana de tipo II invertida) que además de ayudar al ensamblaje de las cadenas  $\alpha/\beta$ , bloquea la hendidura del heterodímero impidiendo la unión del péptido. La unión está mediada por la **proteína BiP** y la **calnexina** (Marks et al., 1995) además de otras chaperonas que ayudan a la estabilización y ensamblaje correcto de complejos nonaméricos ( $\alpha/\beta$ ) 3Ii3 (Roche et al., 1991). Estos complejos son exportados hacia los compartimentos endocíticos (de bajo pH) a través del aparato de Golgi. Una vez en el endosoma, la cadena invariante es digerida por proteasas, quedando en un péptido de unos 20 aa denominado **CLIP** (del inglés, Class II Associated Ii Peptide) que ocupará la hendidura. Aquí, los complejos nonaméricos se disocian en dímeros  $\alpha/\beta$  y entra en juego la molécula HLA-DMA/DMB, que actúa como catalizador en la separación de la molécula HLA de clase II de CLIP, liberando el sitio de unión del péptido para la adquisición de los péptidos antigénicos (Weber et al., 1996), además DM podría desempeñar un papel en la liberación de péptidos distintos de CLIP, seleccionando una determinada subpoblación de péptidos de alta estabilidad para su presentación en la superficie celular. Dicho compartimento endocítico se denomina **MIIC** (del inglés, MHC Class II Compartment) (Gomez-Casado et al., 2006).



## **7. HLA Y FARMACOGENÓMICA.**

Término que hace referencia al estudio de la contribución de los polimorfismos genéticos de un individuo a la variación en las respuestas a los medicamentos entre la población, para asegurar la eficacia terapéutica de los mismos y minimizar los posibles efectos adversos de acuerdo a los perfiles genéticos.

Desde la mejora en la precisión de la determinación de alelos HLA y los fenotipos clínicos en las reacciones a fármacos, han sido publicados un gran número de asociaciones entre los alelos HLA de cada individuo y el desarrollo de varias RAF (Reacciones Adversas a Fármacos, en inglés, *Adverse Drug Reactions*) y SHF (Síndromes de Hipersensibilidad a Fármacos) (Pavlos et al., 2012).

Los residuos clave en la unión de la molécula HLA al péptido de determinados alelos, permiten la formación de enlaces no covalentes con ciertos fármacos. Por lo tanto, la fuerte especificidad entre un alelo HLA y un fármaco puede ser explicada por una complementariedad estérica junto con otras fuertes interacciones no covalentes entre la molécula del fármaco y la hendidura de presentación antigénica, alterando así el repertorio de péptidos ligando que se pueden unir y ser presentados a los linfocitos T (Pavlos et al., 2012).

El concepto de interacción no covalente entre el fármaco y la molécula HLA-TCR puede ocurrir con el fármaco original o con sus metabolitos.

En la última década, se ha producido un incremento significativo del volumen de publicaciones que asocian distintos alelos HLA con diferentes SHF. En general, las diferentes poblaciones tienen distintas asociaciones genéticas. Estas diferencias podrían ser el reflejo de diversas frecuencias alélicas, así como la presencia de otros genes que contribuyen al desarrollo de los distintos síndromes, tales como los genes implicados en el metabolismo de fármacos.(Pavlos et al., 2012).

Ciertos fármacos pueden producir enfermedades como el síndrome de Steven-Johnson/necrólisis epidérmica tóxica (SSJ/NET), que consisten en reacciones mucocutáneas poco frecuentes que destacan por su gravedad.

Los fármacos que causan esta enfermedad, generalmente se superponen con los que causan SHF, siendo los más comunes el alopurinol, anticonvulsivantes aminos aromáticas, antirretrovirales (particularmente Nevirapina), los AINEs y sulfa antimicrobianos (Cheng et al., 2014). En 2004, Chung W.-H. y colaboradores, detectaron una fuerte asociación entre el alelo HLA-B\*15:02 con el SSJ/NET inducido por carbamazepina.(Chung et al., 2004). Observaron que dicho alelo era específico para

la inducción por cabamezepinas de la activación de linfocitos T CD8 en pacientes con SSJ/NET. Por otro lado, también existe una asociación significativa entre el alelo HLA-B\*58:01 y el SSJ/NET inducido por alopurinol en poblaciones orientales, sudasiáticas y caucasoides. (ver tabla II).(Cheng et al., 2014; Pavlos et al., 2012).

Tabla 11: Farmacogenómica de los alelos HLA asociados a reacciones adversas y síndromes de hipersensibilidad a fármacos.

Fármaco	Uso	Efecto adverso	Antígenos o alelos HLA asociados	Población
Abacavir	Inhibidor de la Transcriptasa (VIH-1)	SHF	B*57:01	Africanos, caucosoides
Alopurinol	Gota e Hiperuricemia	SSJ	B*58:01	Orientales, sudasiáticos y caucosoides
Aminopenicilina	Antibiótico	SHF retardada	A2, DR52	Caucosoides
Aspirina	Analgésico	Urticaria	DRB1*13:02; DQB1*06:09; DPB1*02:01	Orientales
		Asma	DPB1*03:01	Orientales
Carbamazepina	Anticonvulsivo	SSJ	A*24:02; A*31:01; B*15:02; B*15:11; B*15:18; B*59:01	Orientales, sudasiáticos y caucosoides
		SHF	A*31:01; A11; B51	Caucosoides y orientales
		Erupción retardada	A*31:01	Caucosoides y orientales
Clozapina	Antipsicótico	Agranulocitosis	B38; DR4; DQ3	Sudasiáticos
Coamoxiclav	Combinación de antibióticos	Daño hepático	DRB1*15:01; DQB1*06:02	Europeos
Efavirenz	Inhibidor de la Transcriptasa (VIH-1)	Erupción retardada	DRB1*01	Caucosoides
Fenitoína	Antiepiléptico	SSJ	B*15:02; B*13:01; C*08:01; DRB1*16:02	Orientales
Feprozona	AINE	Erupción fija por fármacos	B22	Caucosoides
Flucloxacilina	Antibiótico betalactámico	Daño hepático	B*57:01; DRB1*01:07; DQB1*01:03	Caucosoides
Hidralazina	Antihipertensivo	LEIF	DR4	Caucosoides
Lamotrigina	Antiepiléptico	SSJ/NET	B*38:01	Caucosoides
Lapatinib	Quimioterapia	Daño hepático	DRB1*07:01; DQA2*02:01; DQB1*02:02	Internacional
Levanisol	Antihelmínico	Agranulocitosis	B27	Amerindios
Metazolamida	Glaucoma	SSJ	B*59:01; C*01:02	Orientales
Nevirapina	Inhibidor de la Transcriptasa (VIH-1)	Erupción retardada	B*35:05; Cw4; DRB1*01	Caucosoides, africanos y asiáticos
		SHF	B*35:05; DRB1*15; Cw4; Cw8	Asiáticos, caucásicos y orientales
Oxcarbazepina	Antiepiléptico	SSJ	B*15:02; B*15:18	Orientales
		Erupción retardada	B*15:02	Orientales
Oxicam	AINE	SSJ/NET	B*73:01	Caucosoides
D-penicilamina	Proteinuria	Artritis reumatoide	B8; DR3	Caucosoides
Procainamida	Antiarrítmico	LEIF	DR4	Caucosoides
Sulfametoxazol	Antibiótico	SSJ/NET	B*38:02	Caucosoides
Sulfamidas	Antibiótico	SSJ	A29; B12; DR7	Caucosoides
Ximelagatran	Anticoagulante	Daño hepático	DRB1*07:01; DQA1*02	Caucosoides

SSJ: Síndrome de Stevens-Johnson; SHF: Síndrome de Hipersensibilidad a Fármacos; LEIF: Lupus Eritematoso Inducido por Fármacos; NET: Necrolisis Epidérmica Tóxica; AINE: Antiinflamatorio No Esteroideo (Becquemont, 2010; Pavlos et al., 2012)

## **8. HLA Y TRASPLANTE.**

Estudios realizados tras seguimientos de pacientes con órganos transplantados, han comprobado que la compatibilidad entre donante y receptor en las moléculas HLA es necesaria para la supervivencia del injerto en los casos de riñón, páncreas, pulmón, corazón, córnea, hígado e intestino. Esta compatibilidad, debe ser, al menos, con los antígenos pertenecientes a los loci HLA-B y -DR. La compatibilidad debe de ser máxima cuando hablamos de trasplante de médula ósea para evitar que se produzca el llamado rechazo de “injerto contra huésped”; un único cambio en un aminoácido es suficiente para que se produzca este rechazo.

Si se analiza el tiempo de supervivencia de un injerto, se demuestra que cuanto mayor es la compatibilidad HLA, mayor es la supervivencia del mismo. Por ello, es necesario realizar la tipificación HLA de aquellos pacientes que entran en lista de espera para recibir un determinado órgano.

## **9. HLA Y ENFERMEDAD.**

Existen muchos estudios que relacionan ciertos alelos HLA, como factores de protección y susceptibilidad a padecer ciertas enfermedades. Aún no se sabe a ciencia cierta el mecanismo íntimo que provoca estas asociaciones, únicamente se han observado con valores estadísticos. El primero en observar este tipo de asociaciones fue J. Amiel en 1967, quien describió como el alelo HLA-A2 estaba ligado a la enfermedad de Hodgkin (Aimel, 1967).

Algunas enfermedades humanas ocurren con mayor frecuencia en individuos portadores de determinados alelos. Por ejemplo, la mayor asociación HLA-enfermedad conocida, la presenta la espondilitis anquilosante con el HLA-B27. Sin embargo, estas asociaciones estadísticas no son absolutas, lo cual es indicador de que algunos otros genes, así como factores ambientales, desempeñan su papel en el mecanismo fisiopatogénico de las mismas.

Para entender estas asociaciones, se han postulado varias teorías:

El mimetismo molecular entre ciertos patógenos y péptidos autólogos podría desencadenar una respuesta específica autoinmune.

También se postula que, durante el desarrollo del tejido linfoide en la etapa embrionaria, los linfocitos T deben realizar un aprendizaje para reconocer los antígenos propios de cada organismo. Puede ocurrir que por un fallo en esta maduración (debido a

un aislamiento de ciertas células), existan células que no hayan sido expuestas a las moléculas HLA, y al ser expuestas, estas no las reconozcan como propias y desencadenen una respuesta inmunológica.

Existe también, la teoría de la molécula HLA como receptor de virus, como ocurre en el caso del adenovirus tipo II que se une a moléculas HLA clase I.

Por último, mencionar la hipótesis de la alteración de lo propio, donde el agente que causa la enfermedad alteraría las moléculas del propio organismo, y éstas serían el blanco de las respuestas autoinmunes (Moscoso, 2006).

Tabla I2. Algunas enfermedades asociadas a antígenos HLA.

Enfermedad	Asociación HLA
Artritis reumatoide	DR3/DR4
Lupus eritematoso sistémico	DR2/DR3
Espondilitis anquilosante	B27
Diabetes mellitus tipo I	DR3/DR4
Uveitis	B27
Pénfigo vulgar	DR4

## **II. HLA-G: MOLÉCULA Y GEN.**

Como anteriormente ha sido mencionado, HLA-G se clasifica como una molécula de HLA clase I, no clásica. Se encuentra localizada en el brazo corto del cromosoma 6, región 6p21.3, entre el gen HLA-A y HLA-F (Teixeira et al., 2013).

En 1987, Geraghty y sus colaboradores, clonaron por primera vez el gen de HLA-G, en aquel entonces denominado HLA-6.0, dado que era el número de pares de kilobases que pesaba (Geraghty et al., 1987).

Ya desde un inicio se detectó su restringida expresión a nivel tisular (Shukla et al., 1990) y su homología a los genes HLA-A, HLA-B y HLA-C. Sin embargo, se observó que la proteína difería en los productos de los genes HLA-A, HLA-B y HLA-C, en la falta de un segmento mayoritario de la parte intracelular (Geraghty et al., 1987).

Inicialmente fue descubierta en el citotrofoblasto extraveloso del interfaz materno-fetal, lo que indujo a elaborar teorías acerca del papel que ejercía en los mecanismos de tolerancia (Geraghty et al., 1987; Rebmann et al., 2014).

A partir de ese momento, esta apasionante molécula, fue motivo de numerosos estudios de investigación, que aún hoy en día siguen aportando datos de sus muy diferentes papeles en diversos campos.

### **1. ESTRUCTURA.**

#### **1.1 Estructura génica.**

La estructura génica del gen HLA-G consta de una región promotora 5', 8 exones, 7 intrones y, por último, la región 3'UTR. Su organización intrón/exón es muy similar a los genes clásicos (HLA-A, -B y -C) (Teixeira et al., 2013)

Contiene 4.396 pb. La región promotora 5' contiene una secuencia TATAAA variante, que es TCTAAA. Veinticuatro pares de bases aguas arriba de TCTAAA, se encuentra la secuencia CCAAT. Ambos elementos están conservados en los genes HLA-A y -B (Geraghty et al., 1987).

Existen, a su vez, dos secuencias conservadas en los genes HLA clásicos de clase I, en la región promotora 5', una secuencia palindrómica de 13 pb y una segunda secuencia de 11 pb. Ambas modificadas en el gen HLA-G, la secuencia de 11 pb tiene dos bases sustituidas y la secuencia de 13 pb tiene una delección final de 5 pb, siendo las dos no funcionales en HLA-G (ver tabla I3) (Geraghty et al., 1987).

**Tabla I3:** Secuencias conservadas de ADN localizadas en la región promotora de los genes HLA de clase I (Geraghty et al., 1987). En rojo los cambios producidos en HLA-G.

	Secuencia A	Secuencia B
<b>Consenso HLA</b>	5'TGGGGATTCCCCA3'	5'CATTGGGTGTC3'
<b>HLA-G</b>	TGGGGATT.....	CATTAGGTGAC

Por otro lado, en la zona de exones/intrones del gen HLA-G, también existe alguna variación que mencionar: el exón 7 siempre está ausente del ARNm y debido a la aparición de un codón de terminación localizado en el segundo codón del exón 6 y no se traducen el exón 8 (Donadi et al., 2011).

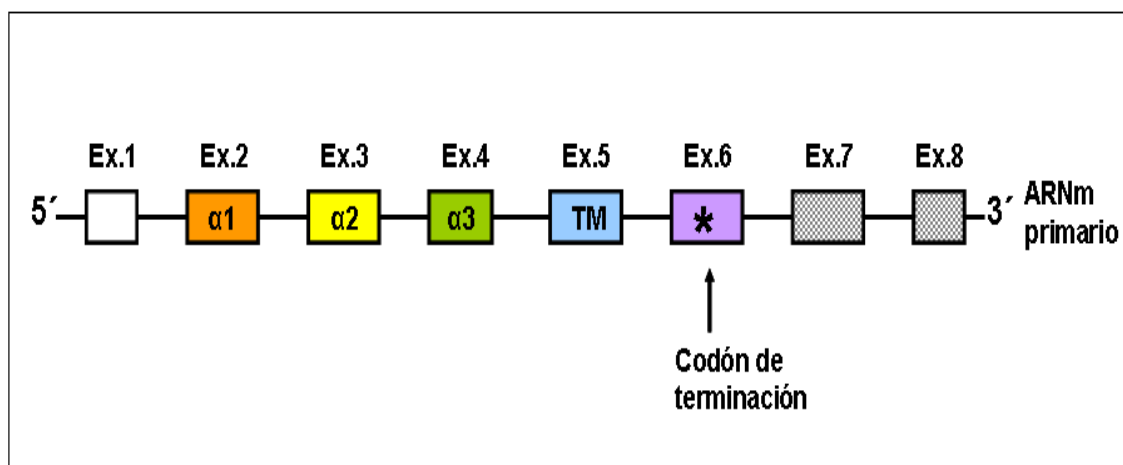


Figura I9: Estructura del gen HLA-G. El exón 7 siempre está ausente del ARNm primario y debido a un codón de terminación (asterisco) en el exón 6, el exón 8 no es traducido. (Ex: exón; TM: Región transmembrana) (Donadi et al., 2011).

## 1.2 Estructura proteica.

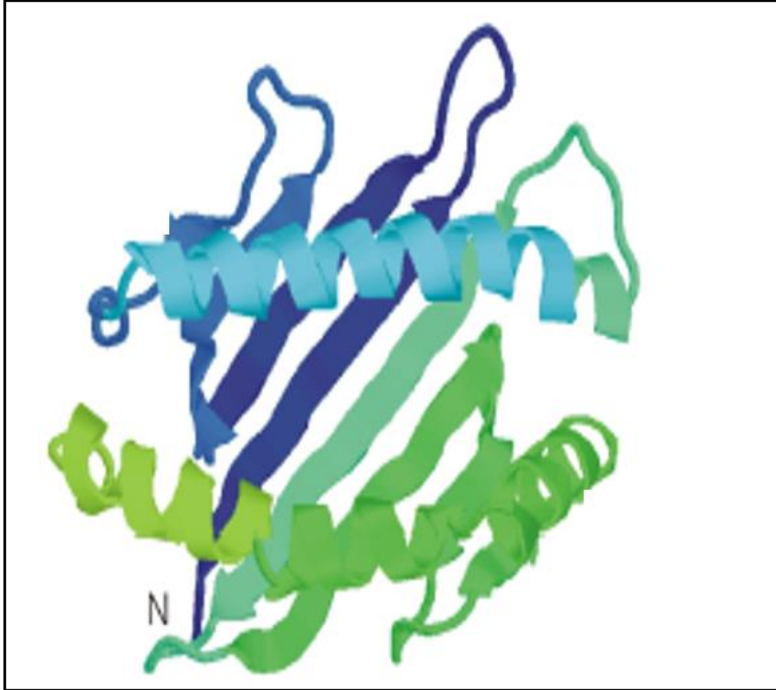
Como se ha visto en apartados anteriores, esta molécula es un heterodímero formado por la cadena pesada  $\alpha$  y la  $\beta$ 2-m. La cadena pesada  $\alpha$  está anclada a la membrana.

La cadena  $\alpha$  consta de los dominios globulares,  $\alpha$ 1,  $\alpha$ 2 y  $\alpha$ 3, idénticos en tamaño a los mismos dominios de las proteínas HLA-A, -B y -C (Gomez-Casado et al., 2006).

En la posición 86 de la región  $\alpha$ 1 también se localiza el aa asparagina N-glicosilada. En los dominios  $\alpha$ 2 y  $\alpha$ 3 existen dos residuos de cisteínas en cada uno, en las posiciones 101 y 164 y en las posiciones 203 y 259, respectivamente. Estas cisteínas están conservadas en los antígenos MHC de clase I y forman dos puentes disulfuro. Dos



El segmento hidrofóbico transmembrana es idéntico en longitud, 27 residuos, a las



El repertorio de péptidos de unión a HLA-G está restringido en comparación con el



Esta molécula posee splicing alternativo en su transcrito primario, lo que da lugar a siete isoformas proteicas, cuatro de ellas ancladas a la membrana (HLA-G1, -G2, -G3, y -G4) y las tres restantes (HLA-G5, -G6, y -G7) solubles, dado que el segmento transmembrana no existe. (Teixeira et al., 2013). HLA-G1 es la isoforma completa, con una estructura similar a las moléculas, ancladas a membrana de HLA clásicas, asociadas a la  $\beta 2$ -m. La isoforma HLA-G2 carece de dominio  $\alpha 2$  (codificada por el exón 3). HLA-G3 no presenta los dominios  $\alpha 2$  ni  $\alpha 3$  (codificada por el exón 3 y 4). HLA-G4 no presenta el dominio  $\alpha 3$  (codificada por el exón 4). Las isoformas solubles HLA-G5 y -G6 presentan el mismo dominio extraglobular que HLA-G1 y -G2, respectivamente, con un bloqueo en la traducción del dominio transmembrana (exón 5). La región 5' del intrón 4 se traduce hasta la aparición de un codón de terminación, confiriendo a las isoformas HLA-G5 y -G6 una cola de 21 aa implicados en su solubilidad. HLA-G7 solo consta del dominio  $\alpha 1$  ligado a dos aa codificados por el intrón 2. Todos los transcritos alternativos carecen del exón 7 (Donadi et al., 2011).

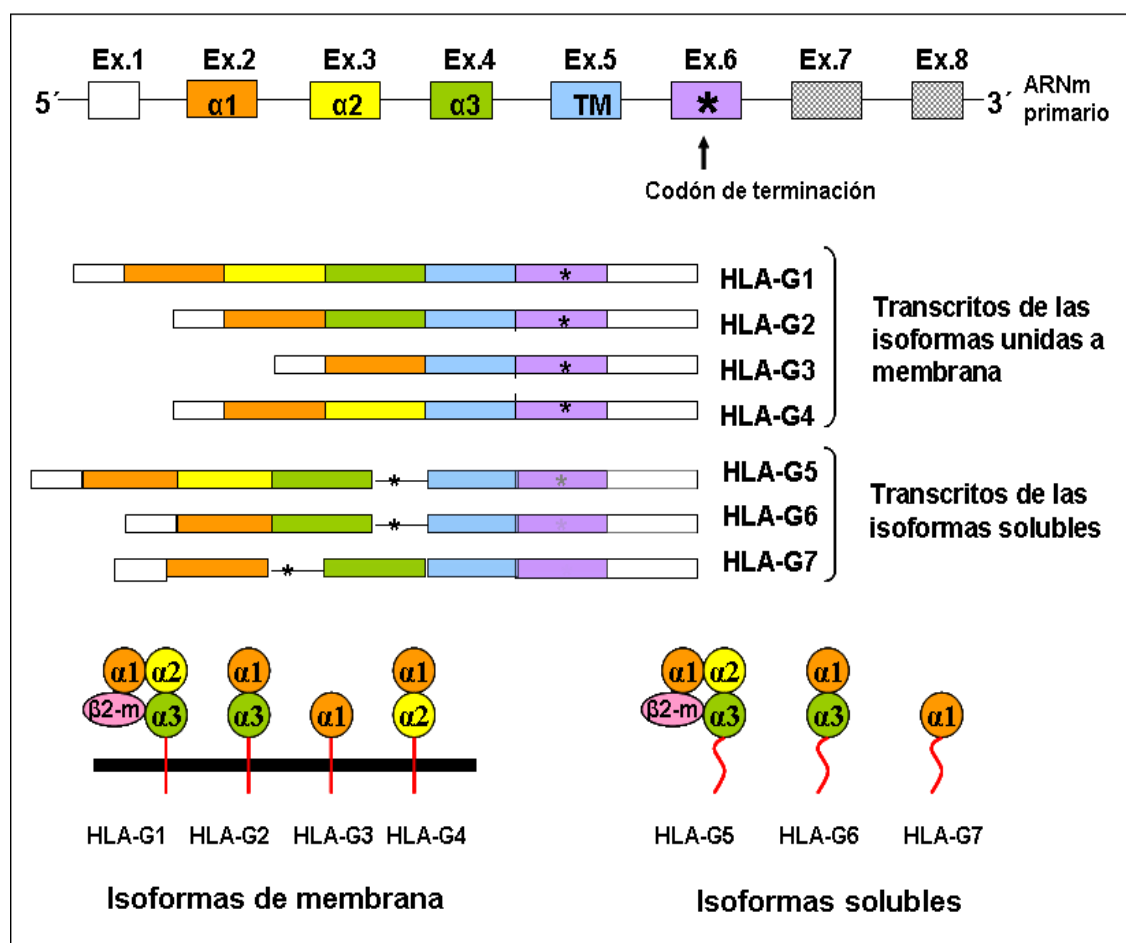


Figura I11: Isoformas de HLA-G producidas por splicing alternativo del ARNm primario. Se esquematizan las siete isoformas de HLA-G (HLA-G1 / HLA-G7). (Donadi et al., 2011).

La mayoría de los alelos actualmente descritos pueden producir teóricamente, por splicing alternativo, todas las isoformas, tanto solubles como ancladas a membrana. Sin embargo, el alelo nulo, HLA-G\*01:05N, presenta la delección de una citosina en el último nucleótido del codón 129 o primer nucleótido del codón 130, en el exón 3, causando un cambio en la pauta de lectura, con la consecuente aparición de un codón de terminación en los siguientes codones: codón 189 (exón 4) que bloquea la traducción de HLA-G1 y -G5, o en el codón 297 en el exón 5, bloqueando la traducción de HLA-G4. Sin embargo, este alelo nulo es capaz de mantener la traducción de las isoformas ancladas a membrana HLA-G2 y -G3 y las solubles HLA-G6 y -G7, ya que en todas ellas el exón 3, que contiene la delección, se ha eliminado por splicing alternativo. (Le Discorde et al., 2005; Suarez et al., 1997).

Existe otro alelo nulo, HLA-G\*01:13N, este presenta la transición de una citosina “C” por una timina “T” en la primera base del codón 54 en el exón 2 (dominio  $\alpha 1$ ), produciendo la formación de un codón de terminación prematuro (TAG) que impide la formación de todas las isoformas, tanto solubles como de membrana, o produce proteínas no funcionales (Lajoie et al., 2008).

### **1.3 Polimerización.**

La presencia de un residuo de cisteína en la posición 42 del dominio  $\alpha 1$ , es una característica única de la molécula HLA-G. Este residuo permite la formación de dímeros gracias a la unión intermolecular Cys42-Cys42, formando un puente disulfuro.

Los dímeros son formados por la recombinación de monómeros de HLA-G, tanto isoformas ancladas a membrana como solubles, mayormente en la superficie celular (Boyson et al., 2002).

Es posible que residuos polimórficos cercanos puedan influenciar la estabilidad del dímero o alterar su flexibilidad en el puente disulfuro. (Donadi et al., 2011).

Una vez formado el dímero, este adquiere una orientación oblicua, exponiendo así, el sitio de unión a CD8 y a LILRB1/B2 (del inglés *leukocyte immunoglobulin-like receptors B1/B2*). Esto permite al dímero, interaccionar con dos LILRB o dos CD8, además de proporcionar mayor afinidad para LILRB en relación a los monómeros. Los dímeros muestran una mayor señalización intracelular mediada por LILRB (Donadi et al., 2011).

Los alelos hasta ahora descritos no presentan ninguna variación en la cisteína del codón 42, por lo que aparentemente, todos los alelos deberían poder formar el dímero.

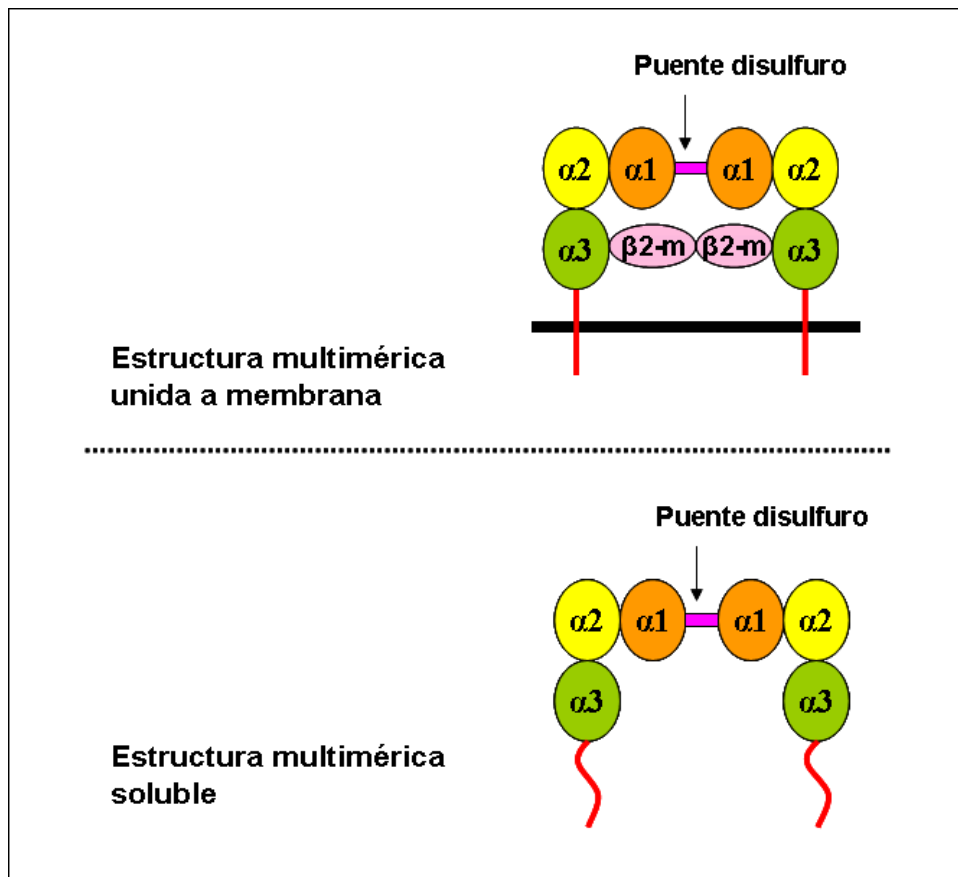


Figura I12: Dímeros HLA-G, anclados a membrana y solubles. (Amodio & Gregori, 2012)

## 2. POLIMORFISMO.

En contraste con el extenso polimorfismo de los genes clase I clásicos de HLA, se ha considerado bajo el grado de polimorfismo en el locus HLA-G, presentando una limitada heterogeneidad de secuencias y con pocos alelos descritos hasta la fecha.

### 2.1 Polimorfismo genético.

El polimorfismo en la región codificante del gen HLA-G es relativamente escaso, aunque se distribuye mayoritariamente entre los exones 2, 3 y 4, así como en intrones. La mayoría de los polimorfismos, o bien no provocan cambios a nivel proteico, o ni siquiera afectan a la estructura secundaria de la cadena pesada.

Por otro lado, existen polimorfismos en la región no codificante, que consiste en la región promotora 5' y en la región 3'UTR, muy importantes en la regulación, expresión del gen y la estabilidad del ARNm.

- Región promotora 5': los polimorfismos en esta región están muy cerca de los elementos regulatorios. Existen ciertas características propias de esta molécula, como la modificación en el elemento regulador *enhancer* A y la delección del elemento ISRE, lo que provoca la insensibilidad del promotor de HLA-G a NF- $\kappa$ B e IFN- $\gamma$ . También existen modificaciones en el sitio  $\alpha$  y en el *enhancer* B. (Donadi et al., 2011; Rebmann et al., 2014).

Hasta la fecha, se han identificado 29 SNP. Entre ellos, algunos de los más significativos desde el punto de vista clínico son: SNP en posición 725, asociado con abortos espontáneos, SNP en posiciones 486 y 964, asociados con asma en niños, rechazo agudo en trasplante de riñón etc. (Lynge et al., 2014; Rebmann et al., 2014).

Antes de entrar en los distintos polimorfismos de la región 3'UTR, es aconsejable aclarar que no existe un consenso en la literatura, en relación a las posiciones de las distintas variaciones de nucleótidos. Como se publica en la base de datos IMGT, el primer nucleótido del primer codón de la secuencia del gen HLA-G lleva una adenina en el exón 1 (ATG) con la ampliación de la secuencia hasta la posición 2.838 del intrón 6, usando la secuencia original descrita por Geraghty y colaboradores. (Geraghty et al., 1987; Robinson et al., 2015). Esta es la pauta que se ha seguido en el presente trabajo.

- Región 3'UTR: esta región también incluye elementos de regulación de la expresión espacio-temporal del ARNm, como motivos ricos en AU etc.

Tres importantes polimorfismos están asociados a la regulación de HLA-G en esta región.

Primero, fragmento 14-pb inserción o delección (14-pb ins/del) (5'ATTTGTTCATGCCT-3'), localizado en el exón 8, asociado con diferencias en la estabilidad del ARNm. La presencia de la inserción del fragmento 14-pb, provoca un sitio de splicing alternativo que genera una delección de 92-pb. (Donadi et al., 2011; Rebmann et al., 2014).

Segundo, SNP en posición +3142 C/G que podría ser un objetivo para ciertos micro ARN, que resulta en la degradación del ARNm de HLA-G. (Donadi et al., 2011; Rebmann et al., 2014).

Tercero, SNP en posición +3187 A/G. La presencia del alelo +3187A está relacionado con la disminución de la expresión de HLA-G por su proximidad al motivo rico en AU. (Rebmann et al., 2014).

Estos tres sitios polimórficos asociados a la producción de HLA-G, también pueden estar asociados entre sí. Es remarcable, que la presencia de 14-pb ins está siempre

acompañado del alelo +3142G y +3187A, ambos previamente asociados a una baja disponibilidad de ARNm, lo que indica que dicha disminución de ARNm asociado a 14-pb ins, podría ser consecuencia de la presencia de estos dos últimos polimorfismos (Donadi et al., 2011).

El descubrimiento y aplicación de la PCR (Saiki et al., 1988), ha permitido desarrollar nuevas técnicas moleculares de caracterización del polimorfismo genético de los alelos HLA de clase I y clase II: amplificación e hibridación con oligosondas específicas de alelo o grupos de alelos y la secuenciación de ADN genómico y ADNc.

En concreto, los alelos descritos para HLA-G se pueden ver en las siguientes tablas, aunque ha de aclararse que, de los 53 alelos descritos y nombrados, solo 50 de ellos tienen su secuencia publicada abiertamente, los restantes aún tienen sus secuencias confidenciales, por ello en la tabla I5 solo hay 50 alelos mencionados:

Tabla I4: Datos de los alelos, proteínas y alelos nulos para el gen HLA-G en la actualidad. (Robinson et al., 2015).

Gen	Alelos	Proteínas	Alelos nulos
HLA-G	53	18	2

Tabla I5: Listado de los 50 alelos HLA-G cuya secuencia está publicada en la actualidad. Con un recuadro el alelo descrito en nuestro laboratorio. Subrayado en rojo se marcan los dos alelos nulos. (Robinson et al., 2015)

HLA-G*01:01:01:01	HLA-G*01:01:12	<u>HLA-G*01:05N</u>
HLA-G*01:01:01:02	HLA-G*01:01:13	HLA-G*01:06
HLA-G*01:01:01:03	HLA-G*01:01:14	HLA-G*01:07
HLA-G*01:01:01:04	HLA-G*01:01:15	HLA-G*01:08:01
HLA-G*01:01:01:05	HLA-G*01:01:16	<b>HLA-G*01:08:02</b>
HLA-G*01:01:01:06	HLA-G*01:01:17	HLA-G*01:09
HLA-G*01:01:02:01	HLA-G*01:01:18	HLA-G*01:10
HLA-G*01:01:02:02	HLA-G*01:01:19	HLA-G*01:11
HLA-G*01:01:03:01	HLA-G*01:01:20	HLA-G*01:12
HLA-G*01:01:03:02	HLA-G*01:01:21	<u>HLA-G*01:13N</u>
HLA-G*01:01:03:03	HLA-G*01:02	HLA-G*01:14
HLA-G*01:01:04	HLA-G*01:03:01:01	HLA-G*01:15
HLA-G*01:01:05	HLA-G*01:03:01:02	HLA-G*01:16
HLA-G*01:01:06	HLA-G*01:04:01	HLA-G*01:17
HLA-G*01:01:07	HLA-G*01:04:02	
HLA-G*01:01:08	HLA-G*01:04:03	
HLA-G*01:01:09	HLA-G*01:04:04	
HLA-G*01:01:11	HLA-G*01:04:05	

Tabla I6: Listado de las 18 proteínas HLA-G descritas en la actualidad. Con un recuadro la proteína descrita en nuestro laboratorio. (Robinson et al., 2015).

HLA-G*01:01	HLA-G*01:11
HLA-G*01:02	HLA-G*01:12
HLA-G*01:03	HLA-G*01:14
HLA-G*01:04	HLA-G*01:15
HLA-G*01:06	HLA-G*01:16
HLA-G*01:07	HLA-G*01:17
HLA-G*01:08	HLA-G*01:18
HLA-G*01:09	HLA-G*01:19
HLA-G*01:10	HLA-G*01:20

### 3. FUNCIÓN.

#### 3.1 Modulación de la respuesta inmune.

Muchas líneas de investigación soportan el papel de HLA-G como una molécula tolerogénica, con una gran importancia en la supresión de la respuesta inmune.

La corta cola citoplasmática retiene a la molécula de HLA-G durante más tiempo en el retículo endoplasmático y prolonga su vida media en la superficie celular por la falta de un motivo de endocitosis.

HLA-G actúa a través de tres receptores de inhibición: LILR (B1 y B2), KIR2DL4 (del inglés, *killer cell immunoglobulin-like receptor*, receptor de células NK) y el receptor CD8 (Shiroishi et al., 2006).

El gen KIR codifica para moléculas que exhiben dos (KIR2D) o tres (KIR3D) dominios de inmunoglobulina extracelulares, y pueden tener motivos ITIM (del inglés *immunoreceptor tyrosine-based inhibitory motif*) (L, para cola larga) o no (S, para cola corta) en el dominio citoplasmático. Los genes KIR con la designación “L” se predice que son receptores inhibitorios, mientras que los que se designan con la letra “S” se predice que son receptores activadores (Donadi et al., 2011).

En general, el dominio  $\alpha 1$  de las moléculas HLA de clase I, es un importante sitio de reconocimiento para los receptores KIR. Sin embargo, el sitio de unión de KIR2DL4 a HLA-G no está bien establecido, aunque se cree que podría estar relacionado con los residuos de metionina (Met) en posición 76 (Met 76) y glutamina (Gln) en posición 79, únicos y conservados en el dominio  $\alpha 1$  de HLA-G.

Los receptores LILRB1 se expresan en la superficie de células NK y linfomononucleares, mientras que LILRB2, se expresa principalmente en la superficie de un grupo restringido de células como monocitos y células dendríticas. Ambos poseen motivos ITIM en sus colas citoplasmáticas, desencadenando señales de inhibición. Estos receptores se unen al dominio  $\alpha 3$  y a la  $\beta 2$ -m de la molécula HLA-G. Sin embargo, LILRB2 se une con una mayor afinidad que LILRB1. Concretamente, los residuos de tirosina (Tyr) 36 y arginina (Arg) 38 del LILRB2 se unen al loop 195-197 del dominio  $\alpha 3$  de HLA-G, mientras que LILRB1 se une con sus residuos de Tyr 38 y Tyr 76 al residuo de fenilalanina (Phe) 195 de HLA-G (Shiroishi et al., 2006). Los polimorfismos observados en este dominio podrían influir en su interacción con LILRB, modulando la señal inhibitoria intracelular.

El receptor CD8 también interacciona con el dominio  $\alpha 3$  de las moléculas HLA clásicas y no clásicas, de tal manera, el sitio de unión de este receptor y de LILRB a HLA-G se superponen. LILRB inhiben la unión de CD8 a las moléculas HLA, esto indica que la función inhibitoria de los receptores LILRB es dual, desplazando a CD8 y activando los motivos ITIM (Shiroishi et al., 2003).

Las moléculas solubles de HLA-G pueden inducir apoptosis en linfocitos T CD8+, así como en células NK CD8+. En condiciones en las que HLA-G soluble está elevado, incluyendo el embarazo, algunos tumores y aloinjertos con buen pronóstico, este mecanismo representa un efecto inmunomodulador adicional de HLA-G (Donadi et al., 2011).

### **3.2 HLA-G y su importancia en la gestación.**

El útero y la placenta constituyen el único sitio de modulación inmunológica donde el feto es tolerado por el sistema inmunológico de la madre. (Lynge et al., 2014). Sin embargo, hoy en día se sabe que todos los miembros de la futura familia, el embrión, la madre y el padre, contribuyen a establecer una tolerancia inmunológica a través de HLA-G, antes y durante el embarazo (Dahl et al., 2014).



Se especula que la molécula HLA-G esté envuelta en mecanismos de reproducción incluso antes de la concepción, dado que HLA-G se puede detectar en el tracto genital y en la sangre de mujeres no embarazadas, y en fluidos seminales del hombre, concretamente la isoforma soluble HLA-G5. Además, HLA-G también se ha detectado en embriones pre-implantados (Dahl et al., 2014).

Se sabe que HLA-G inhibe a los linfocitos T y células NK de la madre dirigidos contra las células fetales uniéndose a los receptores KIR2DL4, CD8 y LILRB anteriormente comentados (Rebmann et al., 2014).

Durante la implantación, el desarrollo de transcritos específicos de HLA-G, concretamente las isoformas HLA-G3 y -G4, van aumentando según avanza esta etapa. La isoforma más expresada en los trofoblastos es HLA-G1. La concentración de isoformas solubles, derivada posiblemente de la placenta por escisión de HLA-G1 y la secreción de HLA-G5, aumenta de dos a cinco veces en mujeres embarazadas, contribuyendo a la modulación de citoquinas maternas, en concreto, la regulación a la alta de citoquinas anti-inflamatorias como IL-3, IL-4 y IL-10 (Interleuquinas). (Lynge et al., 2014; Rebmann et al., 2014).

Los niveles de HLA-G soluble se detectan en la sangre periférica de la madre, con el pico más alto en el tercer mes de gestación y comienzan a disminuir durante el tercer trimestre. Cuando estos niveles son indetectables o con variaciones, especialmente en las primeras semanas de gestación, se relaciona con complicaciones en el embarazo.

La pre-eclampsia es un desorden multisistémico del embarazo que se manifiesta clínicamente a finales del segundo y tercer trimestre, con hipertensión acompañada de proteinuria. Las incidencias de este desorden han aumentado un 25% en las últimas dos décadas (Moussa et al., 2014). La etiología es desconocida, aunque ciertos estudios apoyan la teoría basada en una mala adaptación del sistema inmune, con bajos niveles de células reguladoras y baja expresión de HLA-G (Lynge et al., 2014).

El polimorfismo 14-pb ins/del está asociado con pre-eclampsia grave en numerosos estudios. El genotipo 14-pb del, en homocigosis, está relacionado con los niveles más altos de HLA-G soluble, mientras que 14-pb ins, en homocigosis, con los más bajos. (Dahl et al., 2014). Por el contrario, otros estudios no han visto ninguna asociación entre este polimorfismo y la pre-eclampsia.

Existen otros polimorfismos asociados. El alelo HLA-G\*01:06 paterno, se ha visto que contribuye al aumento de riesgo de la pre-eclampsia en mujeres multigrávidas que

no llevan este alelo (Lynge et al., 2014). El alelo +3187A, en la región 3'UTR, también se asocia a esta enfermedad.

Otra alteración importante en el embarazo son los abortos espontáneos y recurrentes. Se ha estudiado la posible asociación entre 14-pb ins y el aumento de riesgo de aborto recurrente, sin embargo otros autores lo impugnan (Lynge et al., 2014).

### **3.3 HLA-G y su papel en trasplante.**

La primera evidencia de que la expresión de HLA-G estaba relacionada con el trasplante fue descrita en el trasplante de corazón. (Rebmann et al., 2014).

En general, respecto al polimorfismo 14-pb ins/del, el genotipo 14-pb ins, en homocigosis, está asociado con el rechazo o la enfermedad de injerto contra huésped (EICH), mientras que el genotipo 14-pb del, en homocigosis, está asociado a una menor probabilidad de rechazo (Donadi et al., 2011).

En el trasplante de hígado, no se observa una diferencia significativa respecto a la frecuencia de los alelos 14-pb ins/del en enfermos, comparados con controles sanos. Sin embargo, la frecuencia de 14-pb ins en homocigosis, aumenta en pacientes con rechazo agudo. (Donadi et al., 2011).

Respecto al trasplante de medula ósea para talasemia, el genotipo 14-pb del, en homocigosis, está asociado a una aguda EICH, mientras que la inserción en homocigosis y heterocigosis con bajo riesgo de EICH.

En relación al trasplante de células progenitoras hematopoyéticas (TCPH), el genotipo 14-pb ins, en homocigosis, está asociado con supervivencia libre de enfermedad. Sin embargo, este genotipo presenta factor de riesgo para EICH después de TCPH, usando la médula ósea de donante HLA compatible. Por otro lado, existen estudios que no detectan diferencias significativas respecto a este polimorfismo y TCPH (Donadi et al., 2011; Rebmann et al., 2014).

Contemplando el trasplante de corazón, el genotipo 14-pb del, en homocigosis, está relacionado con niveles más altos de HLA-G soluble y una mayor biodisponibilidad de ciclosporina (Donadi et al., 2011).

En trasplantes de pulmón, hígado y riñón, el aumento de la expresión de HLA-G está asociado con una mejor aceptación del injerto en términos de rechazo agudo o crónico. La presencia del alelo +3187A, se asocia con el rechazo agudo o crónico después del trasplante de hígado (Rebmann et al., 2014).

Bajo la asunción de que la expresión de HLA-G es de manera preferente controlada por factores genéticos, en trasplante la genética de ambos, el donante y el receptor, se tienen que tener en cuenta. Aunque HLA-G no sea una molécula presentadora de antígenos, la presencia de compatibilidad HLA-G está asociada con menor riesgo de rechazo. (Rebmann et al., 2014).

### **3.4 HLA-G y su doble función en cáncer.**

Las células cancerosas presentan antígenos asociados a tumores, codificadas por genes desregulados mutantes o normales que, una vez presentados por moléculas HLA de clase I clásicas, pueden ser reconocidos por el sistema inmune del huésped, siendo generalmente eliminados. La mayor vía de escape, es la regulación a la baja o falta de expresión de moléculas HLA de clase I clásicas, así las células NK son las que eliminarían los tumores (Donadi et al., 2011).

El concepto de cáncer incluye tres fases: eliminación, equilibrio y evasión. HLA-G puede interferir en cada una de estas fases. Puede regular a la baja la fase de eliminación por inhibición de la proliferación de linfocitos T, B, células NK, la actividad fagocítica de neutrófilos y la función de las células dendríticas vía LILRB. También puede afectar a la fase de equilibrio controlando la expresión de las moléculas HLA de clase II por células dendríticas. Y respecto a la fase de evasión, el tumor en esta fase, pierde moléculas importantes para el reconocimiento inmunológico y tiende a expresar HLA-G en la superficie celular, haciéndolas menos susceptibles a las células efectoras. En esta última fase también aumenta la expresión de citoquinas inmunosupresoras como IL-10. Todo esto resulta en el crecimiento del tumor, invasión y metástasis (Rouas-Freiss et al., 2014).

HLA-G se encuentra en tumores sólidos y metástasis en estadios avanzados, por lo que, ha sido propuesto el uso de HLA-G como marcador de pronóstico, y se ha considerado que los niveles altos de HLA-G serían indicativos de un mal pronóstico (Rouas-Freiss et al., 2014). Así ha sido descrito en individuos con cáncer colorrectal, gástrico, de pulmón o de esófago (Cao et al., 2011).

Se ha observado que, en tumores hematológicos, las células tumorales son células inmunes capaces de expresar receptores LILRB1, por ello, esta molécula podría tomar un papel inesperado en la inhibición de la proliferación celular. HLA-G inhibiría la proliferación de linfomas de células B, mielomas y leucemias de células B ya que

expresan LILRB1. Esto supone una acción de HLA-G totalmente opuesta a la explicada anteriormente en tumores no hematológicos. (Naji et al., 2012).

En relación al polimorfismo 14-bp ins/del de la región 3'UTR, Ge YZ y colaboradores detectaron, que debe contribuir a la susceptibilidad al cáncer de mama y en general al riesgo de cáncer en la población asiática. (Ge et al., 2014).

El carcinoma hepatocelular (CHC) es un tumor maligno primario del hígado. El alelo 14-pb del, es más frecuente en casos de CHC en la población brasileña. HLA-G soluble está más elevado en el plasma de estos pacientes. HLA-G\*01:05N, está asociado con hepatitis B o hepatitis B con CHC, jugando un papel importante en la supervivencia inmunitaria en la hepatitis B y CHC (Teixeira et al., 2013).

Respecto al cáncer de mama, la presencia de 14-pb, está asociada a mayor susceptibilidad a este cáncer. Los niveles de HLA-G soluble en circulación están aumentados. (Jeong et al., 2014).

En el cáncer de vejiga, el locus HLA-G está asociado a mayor susceptibilidad al desarrollo y progresión del carcinoma de células transicionales (CCT). La familia de alelos HLA-G\*01:04, especialmente -G\*01:04:04, y -G\*01:03, están asociados a la dependencia del tabaco con desarrollo a CCT. El alelo HLA-G\*01:04 se asocia al peor pronóstico de este cáncer, mientras que -G\*01:03 se asocia a una protección contra el CCT. (Donadi et al., 2011).

En la infección por el virus del papiloma humano (VPH), los alelos HLA-G\*01:03 y -G\*01:04 se asocian con lesión intraepitelial escamosa (LIE), y el genotipo G\*01:01/G\*01:04 con LIE de alto grado. (Donadi et al., 2011).

### **3.5 HLA-G y su relación con enfermedades autoinmunes.**

La autoinmunidad es causada por un deterioro de la tolerancia. Se sabe que HLA-G es una molécula tolerogénica y que la presencia de HLA-G soluble o su expresión en tejidos afectados, debería disminuir las manifestaciones autoinmunes. Por ello, HLA-G tendría que ser un elemento ventajoso para la enfermedad, y por lo tanto, estudiado en este aspecto, aunque no ha sido así. (Donadi et al., 2011).

En la esclerosis sistémica, la expresión de HLA-G se asocia con una menor frecuencia de úlceras vasculares cutáneas, telangiectasia, poliartritis, y con mayor supervivencia (Brenol et al., 2012).

Un reciente meta-análisis demuestra que el polimorfismo 14-bp ins/del se asocia con la susceptibilidad a lupus eritematoso sistémico (LES), se observa un aumento en la

frecuencia del alelo 14-pb ins, así como el genotipo en homocigosis (Donadi et al., 2011). También el polimorfismo +3142 G/C está asociado con la susceptibilidad a LES en población suramericana. (Lee et al., 2014).

En relación a la artritis reumatoide (AR), el alelo 14-pb del, está presente en pacientes con AR en su forma juvenil pero no en la adulta. Por otro lado, el genotipo 14-pb ins en homocigosis, se asocia a una mejor respuesta terapéutica a metotrexato (fármaco contra la AR) en pacientes adultos (Donadi et al., 2011). Otros estudios no observan asociación ninguna. (Lee et al., 2014)

Se cree que HLA-G actúa como una molécula anti-inflamatoria en esclerosis múltiple (EM). Se ha observado que los niveles de HLA-G soluble en suero y líquido cefalorraquídeo, pueden estar influenciados por polimorfismos HLA-G, independientemente de la inflamación. (Rizzo et al., 2012). Otras investigaciones, no han encontrado asociación en las frecuencias de los polimorfismos -725C/G (de la región promotora), las posiciones 129 y 130 del exón 3, 14-pb ins/del y EM (Donadi et al., 2011).

Respecto a la psoriasis, se ha sugerido que HLA-G podría prevenir la destrucción de queratinocitos por regulación de la actividad de linfocitos citotóxicos (Sweeney & Kirby, 2011). También se ha detectado cierta ventaja en el alelo HLA-G 14-pb del y el genotipo en homocigosis, ya que se asocian con una mejor respuesta a acitretina (fármaco antipsoriático) (Borghi et al., 2014).

La asociación del asma bronquial con HLA-G, indica que el genotipo GG del polimorfismo -964G/A, se asocia con asma en niños con madres afectadas, mientras que el genotipo AA se asocia con niños con madres no afectadas (Donadi et al., 2011).

El alelo 14-pb ins, se ha observado en alta frecuencia en pacientes con sarcoidosis y la enfermedad de Behcet, mientras que el alelo 14-pb del, se ha publicado como un factor de riesgo para el pénfigo vulgar y la miocardiopatía dilatada. (Brenol et al., 2012).

### **3.6 HLA-G y su relación con el VIH.**

El síndrome de la inmunodeficiencia adquirida (SIDA) fue mundialmente conocido después de los años 80 cuando aparecieron las primeras señales de epidemia, y hasta ahora no hay cura o vacuna efectiva para esta enfermedad (da Silva et al., 2014).

A través de ciertos mecanismos, los virus han desarrollado la capacidad de impedir el reconocimiento de las células NK, mediante la inducción de la expresión de HLA-G.

Las propiedades inmunosupresoras de HLA-G podrían contribuir a la susceptibilidad y persistencia de las infecciones virales (Donadi et al., 2011). Existen dos tipos de VIH, VIH-1 y VIH-2 (da Silva et al., 2014).

Se deben tener en cuenta los polimorfismos HLA-G. El alelo nulo HLA-G\*01:05N, está asociado con protección frente a la infección por VIH-1, mientras que el alelo HLA-G\*01:01:08, se asocia a un mayor riesgo de infección en mujeres de Zimbawe. Dado que el alelo HLA-G\*01:05N no codifica para la proteína completa HLA-G1 y como esta es el mayor ligando por el cual HLA-G inhibe a las células NK, se ha postulado que la ausencia o reducción de la expresión de HLA-G1, podría permitir a las NK destruir las células infectadas por el VIH, dando protección a la infección. (Matte et al., 2004).

Mundialmente, 1,4 millones de mujeres embarazadas e infectadas con el virus, pueden transmitir el VIH a sus hijos durante la gestación, el parto o después del nacimiento mediante la lactancia materna (Segat et al., 2014). La transmisión vertical del VIH-1 es la ruta primaria para la infección de los hijos (Luo et al., 2013).

Se desconoce por qué algunos niños son infectados y otros no, a pesar de una exposición significativa al VIH. Existen ciertos factores independientes, asociados con la transmisión vertical que son: cargas elevadas del virus en plasma, baja cantidad de linfocitos T CD4+ en la madre, parto vaginal, baja edad gestacional, no uso de fármacos antirretrovirales por la madre y/o el neonato, la lactancia materna y HLA-G expresado en los citotrofoblastos de la interfase materno-fetal (Luo et al., 2013; Segat et al., 2014).

Se ha observado que las madres con el alelo HLA-G\*01:03 tiene menos probabilidad de transmitir el virus a sus hijos durante el periodo perinatal. Podría ser debido a las diferentes variantes expresadas de HLA-G y/o a las interacciones con los KIR, aunque requeriría más estudio. La presencia de los alelos HLA-G\*01:01:01:03 y HLA-G\*01:05N también está asociada a un menor riesgo de transmisión vertical del VIH-1 (Segat et al., 2014).

La presencia del alelo 14-pb del y del genotipo en homocigosis, están asociados con un efecto protector respecto a infección perinatal de VIH-1 en niños brasileños y en americanos africanos, sugiriendo que la mayor expresión de HLA-G podría contribuir a crear una respuesta inmune local Th2, disminuyendo la producción de citoquinas pro-inflamatorias asociadas a un aumento del riesgo de transmisión vertical. Por el contrario, esta asociación no se ha detectado en población europea. Por otro lado, la presencia del alelo 14-pb ins está asociado con protección frente a la infección en niños

de Zambia (Segat et al., 2014). El alelo +3010C por separado, no tiene efecto en la transmisión vertical del VIH-1, sin embargo, al encontrarse asociado a 14-pb del (ambos en desequilibrio de ligamiento), ejerce un papel positivo en la protección frente al virus. (Donadi et al., 2011).

Así, las distintas variantes genéticas de HLA-G podrían afectar en distinta medida a la susceptibilidad frente a la infección, teniendo en cuenta la población de la que se trate. Por lo que, sería necesario una mayor investigación en este campo.

### **III. POBLACIÓN DE ESTUDIO: AMERINDIOS INMIGRANTES EN MADRID.**

#### **1. EL ORIGEN DE LOS AMERINDIOS.**

La población americana, está constituida por tres grupos principales de habitantes. El primero lo componen los descendientes de las poblaciones indígenas, es decir, individuos que no se mezclaron con la población que llegó desde Europa durante el descubrimiento y posterior asentamiento de tropas en América en el siglo XV y XVI. Estas poblaciones viven mayoritariamente en zonas rurales y organizadas en pequeños grupos por toda la geografía. El segundo grupo se encuentra constituido por la población mestiza (mezcla de nativos con europeos), representa el mayor porcentaje de la población y puede presentar diferentes grados de mestizaje dependiendo de la región, y ocupan principalmente los núcleos urbanos de población. El último grupo es el denominado afro-americano, son principalmente mezcla entre africanos, europeos y nativos.

Estas poblaciones aparecen poco relacionadas con poblaciones del resto del mundo, lo que hace suponer que el origen de estos grupos puede haberse iniciado en el viejo continente y presentar una evolución divergente determinada principalmente por factores medioambientales que llevó a una diferenciación genética específica.

Existen varias teorías acerca del poblamiento de América. Una de las más significativas es la teoría de Greenberg.

En el año 1986, el antropólogo JH Greenberg propuso la teoría de las “tres oleadas migratorias”, apoyada en un estudio multidisciplinar (genético, dental y lingüístico). Dichas oleadas, darían origen a los tres grupos de poblaciones primitivas en el continente americano: amerindios (la mayor parte de los indios de América del norte y de América del sur; 12.000 años AC); Na-Dene (atabascos, navajos y apaches; 8.000 años AC) y Esquimo-Aleutianos (6.000 años AC) (Greenberg et al., 1986).

Por otro lado, existen otras hipótesis como La Ruta Transpacífica, originaria desde Asia o Polinesia. Está basada en resultados de distintos estudios de genética molecular y de enfermedad. Propone la posibilidad de una ruta transpacífica como vía de llegada a América para los primeros pobladores del Nuevo Mundo.

La Ruta Transatlántica, teoría que propone el poblamiento de América desde el continente europeo a través del Océano Atlántico, antes del descubrimiento de América



(Arnaiz-Villena A et al., 2014; Arnaiz-Villena et al., 2000; Arnaiz-Villena et al., 2010). Arqueólogos han encontrado en Nuevo México herramientas utilizadas hace 20.000 años en España (Arnaiz-Villena et al., 2000)

Las poblaciones nativas amerindias, se encuentran distribuidas a través de todo el territorio americano especialmente en Centro y Suramérica. En la región de Norteamérica, el número de poblaciones consideradas como amerindias se encuentran en números bastante reducidos. La familia amerindia contiene 583 lenguas, habladas por 18 millones de individuos.(Arnaiz-Villena et al., 2010)

## **2. ESTUDIO DEL SISTEMA HLA EN POBLACIONES AMERINDIAS.**

Tecnologías como, estudios de polimorfismos de ADN en el cromosoma Y, análisis de ADN mitocondrial y análisis de los genes HLA entre otros, han incrementado favorablemente el conocimiento de las características genéticas de los pobladores de América y permiten su comparación con otras poblaciones alrededor del mundo.

Diferentes estudios HLA en amerindios han mostrado resultados sorprendentes, que hacen que este grupo de poblaciones sean de especial interés y de gran ayuda para poder comprender con mayor precisión algunos aspectos no aclarados todavía, como las posibles rutas migratorias que han dado origen al poblamiento de América, o la evolución del polimorfismo de los genes HLA después del contacto con nuevos patógenos que su sistema inmune fue incapaz de responder (a partir del siglo XV).

### **2.1 Polimorfismos de los genes HLA en amerindios.**

Una característica destacada del sistema HLA en amerindios, es el bajo número de alelos que hay en los genes de clase I y de clase II (Cadavid & Watkins, 1997). El número bajo de alelos presente en los grupos amerindios, se ha relacionado con la hipótesis que considera a la población de América como originada por un número reducido de cazadores provenientes de Siberia con un número bajo de alelos HLA.

#### **Locus HLA-A**

A nivel del locus HLA-A, los grupos de alelos más representativos son: HLA-A\*02, -A\*24, -A\*31 y -A\*68. (Gonzalez-Galarza et al., 2011; Meyer et al., 2007).

#### **Locus HLA-B**

Los grupos de alelos más representativos del locus HLA-B son: HLA-B\*15, -B\*27, -B\*35, -B\*39, -B\*40 y -B\*48 (Gonzalez-Galarza et al., 2011; Meyer et al., 2007).

#### **Locus HLA-C**

En relación al locus HLA-C, las poblaciones amerindias tienen preferentemente los siguientes grupos de alelos: HLA-C\*01, -C\*03, -C\*04, -C\*07, -C\*08 etc. (Gonzalez-Galarza et al., 2011; Meyer et al., 2007).

#### Locus HLA-DRB1 y DQB1

El comportamiento de los loci HLA de clase II en amerindios, es diferente a los pertenecientes a la clase I, ya que no existe una variabilidad tan alta como en los últimos. Sin embargo, en la metodología para estudios de emparentamiento poblacional, está estandarizado el uso de estos loci, en especial HLA-DRB1, dado que para la mayoría de las poblaciones están determinados los alelos de ambos, con cuatro dígitos de resolución (Rey et al., 2013).

Respecto al locus HLA-DRB1, los grupos de alelos más frecuentes son: HLA-DRB1\*04, -DRB1\*08, -DRB1\*14 y -DRB1\*16 (Gonzalez-Galarza et al., 2011; Meyer et al., 2007).

Acerca del locus HLA-DQB1, los grupos de alelos más frecuentes son: HLA-DQB1\*03, -DQB1\*04 y -DQB1\*05 (Gonzalez-Galarza et al., 2011; Meyer et al., 2007).

#### Haplotipos extendidos más frecuentes

a) Haplotipos centroamericanos: son compartidos con otros amerindios y uno de ellos (A\*02-B\*35-DRB1\*08:02-DQB1\*04:02) con la población Yupik (aborígenes esquimales del norte del continente americano). (Arnaiz-Villena et al., 2010).

- HLA-A\*02-B\*35-DRB1\*04:07-DQB1\*03:02
- HLA-A\*02-B\*39-DRB1\*04:07-DQB1\*03:02
- HLA-A\*02-B\*35-DRB1\*08:02-DQB1\*04:02
- HLA-A\*24-B\*39-DRB1\*04:07-DQB1\*03:02
- HLA-A\*68-B\*39-DRB1\*04:07-DQB1\*03:02

b) Haplotipos suramericanos: la mayoría de los haplotipos frecuentes son compartidos con otros amerindios estadounidenses, pero no con asiáticos. (Arnaiz-Villena et al., 2010).

- HLA-A\*02-B\*35-DRB1\*08:02-DQB1\*04:02
- HLA-A\*02-B\*15-DRB1\*08:02-DQB1\*04:02
- HLA-A\*24-B\*35-DRB1\*08:02-DQB1\*04:02
- HLA-A\*02-B\*35-DRB1\*04:03-DQB1\*03:02
- HLA-A\*02-B\*40-DRB1\*04:07-DQB1\*03:02

De tal forma, se observa que los haplotipos extendidos de los amerindios no comparten nada con los haplotipos asiáticos, por lo que genéticamente hablando ambas poblaciones están alejadas (Arnaiz-Villena et al., 2010).

## **2.2 Relaciones genéticas en las poblaciones amerindias.**

Los amerindios mayoritariamente viven en Centro y Suramérica, donde además hay otras poblaciones más numerosas, como las mestizas (mezcladas con europeos) y las negras (mezcladas con poblaciones negroides de África). Este fenómeno es importante, dado que, en una misma área, pueden vivir grupos con diferencias genéticas relacionadas a su vez con patologías específicas, así como, con diferentes modelos de respuesta a enfermedades infecciosas y/o enfermedades autoinmunes. El conocimiento de las características genéticas de cada uno de ellos es de gran utilidad, ya que, permite delinear el posible comportamiento de cada grupo con relación a algunas enfermedades y a su vez, desarrollar programas de medicina preventiva y protocolos de “*screening*” para la detección precoz de patologías que requieren un manejo temprano (Erllich et al., 1993; Perez-Luque et al., 2000).

Los estudios HLA realizados en distintos grupos étnicos, han dado resultados parecidos: a) son grupos poco polimórficos, hay pocas variantes alélicas en cada grupo poblacional; b) estas variantes, presentan frecuencias muy elevadas, llegando hasta más del 50% en muchas poblaciones para ciertos alelos; c) los alelos presentes en alta frecuencia, no lo están en otros grupos poblacionales del resto del mundo, donde ni tan siquiera existen estos alelos; d) algunos haplotipos HLA extendidos encontrados en los amerindios, son comunes a otras poblaciones amerindias y no a poblaciones de otras regiones. (Moscoso, 2006) (Arnaiz-Villena A et al., 2014; Arnaiz-Villena et al., 2000; Arnaiz-Villena et al., 2010)

Cabe mencionar también que: a- los amerindios de mesoamérica (relacionados con las culturas Maya y Olmeca) parecen menos relacionados con las islas del Pacífico que algunos amerindios de sudamérica; b- la Isla de Pascua muestra un probable intercambio cultural y genético con aymaras del lago Titikaka. Esta civilización también comparte rasgos significativos con las poblaciones ibéricas de construcciones megalíticas en Europa; c- los mesoamerindios parecen agruparse de acuerdo a los genes HLA y a rasgos culturales (la cultura Olmeca podría haber dado lugar a los mayas, zapotecos, mixtecos, mazatecos y culturas Mixe); d- todos los amerindios del Norte, Meso y Sudamérica se agrupan juntos y se separan de otras poblaciones mundiales. Las

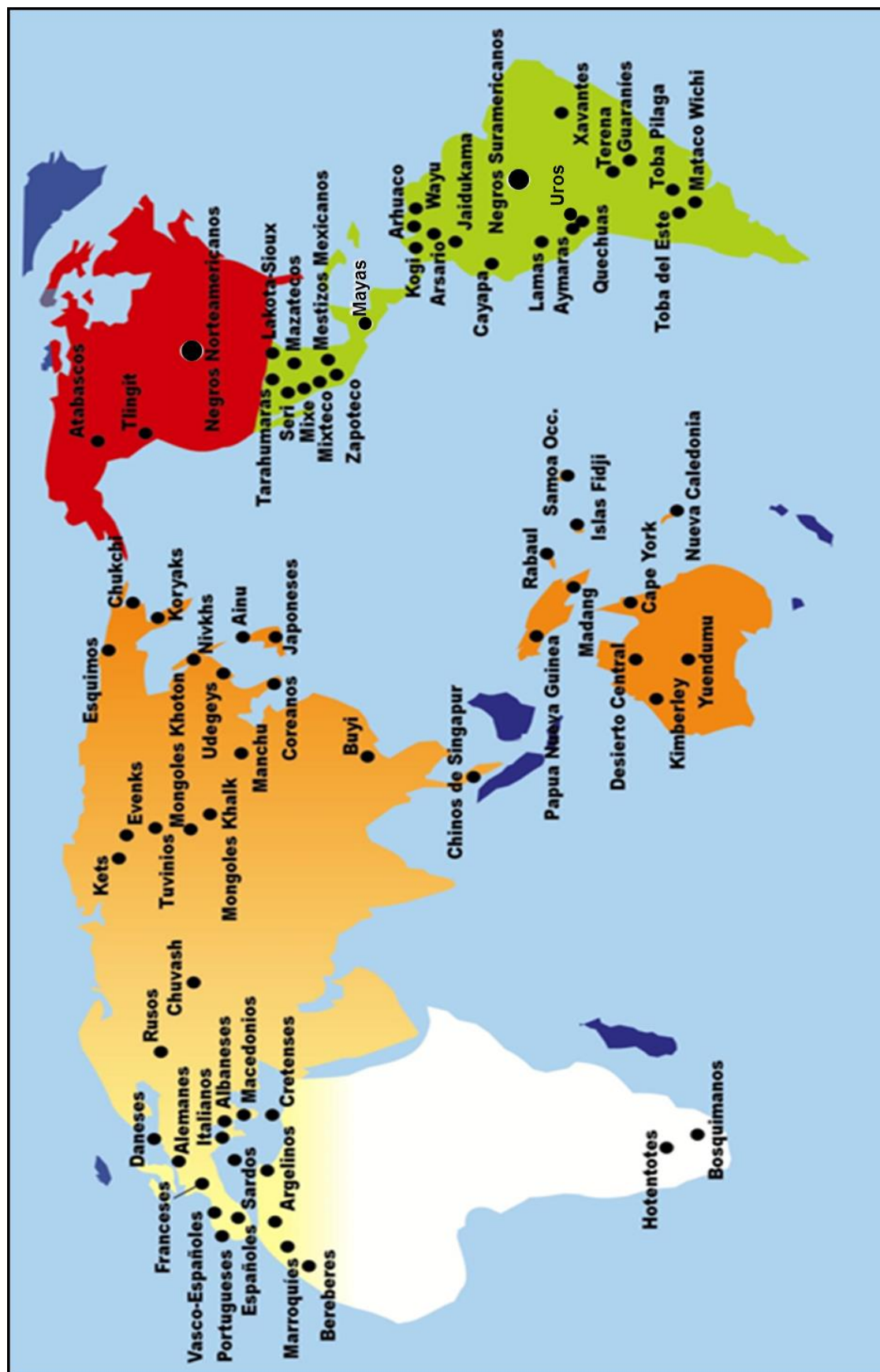
poblaciones amerindias siguen poco desnivel geográfico respecto a las frecuencias HLA analizadas, a excepción de grupos mesoamericanos. (Arnaiz-Villena A et al., 2014; Arnaiz-Villena et al., 2000; Arnaiz-Villena et al., 2010)

Finalmente, los estudios HLA en las poblaciones de América, son una herramienta útil para la caracterización genética de cada uno de los grupos étnicos y esta sería una base para el análisis de los posibles riesgos de patologías infecciosas y autoinmunes, así como también el apoyo fundamental para los programas de transplante de órganos y tejidos.

### **3. AMERINDIOS INMIGRANTES EN MADRID.**

La población amerindia tiene un alto grado de homogeneidad genética entre sí, mientras que hay grandes diferencias antropológicas e inmunogenéticas frente al resto de poblaciones mundiales, incluyendo a la española. (Parga-Lozano et al., 2011).

Mientras todas las poblaciones mundiales están emparentadas genéticamente en cuanto al perfil de genes HLA (en el mapa se representa con un suave gradiente de rojo a amarillo), los amerindios tienen un perfil HLA totalmente diferente y así se observa un cambio brusco de color rojo a verde en la figura I13.



**Figura I13:** Mapa mundial de las poblaciones determinadas para los loci HLA-DRB1 y -DQB1. Las poblaciones amerindias figuran separadas del resto, con un color verde intenso. (Arnaiz-Villena et al., 2010; Arnaiz-Villena et al., 2000; Arnaiz-Villena et al., 2014)

Desde hace más de diez años, el flujo migratorio de amerindios (principalmente andinos) a las grandes ciudades españolas como son Barcelona y Madrid, ha hecho que buena parte del sector servicios, esté ocupado por trabajadores amerindios en estos núcleos urbanos. Alrededor de un 4,0-4,5% de la población de la Comunidad de Madrid era amerindia en el momento de la toma de muestra. (Parga-Lozano et al., 2011)(<http://www.madrid.org/>). Según el barómetro de inmigración en la Comunidad de Madrid, en el 2014, el país amerindio más representado en la comunidad inmigrante fue Ecuador, seguido de Colombia, Perú, Bolivia, República Dominicana, Paraguay y, por último, Brasil. Aunque es remarcable que, en estos últimos años, debido a la situación económica de España, muchos de estos inmigrantes han vuelto a sus lugares de origen, mayoritariamente los ecuatorianos (<http://www.madrid.org/>).

# OBJETIVOS





- 1- Determinar alelos de HLA-G en una población de amerindios inmigrantes en Madrid.
- 2- Calcular los haplotipos extendidos de HLA-G/HLA-A/HLA-DRB1 (clase I y clase II) en la misma población.
- 3- Comparar las frecuencias de alelos HLA-G amerindios con los del resto de poblaciones mundiales.
- 4- Buscar posibles alelos nuevos del locus HLA-G en amerindios.
- 5- Extraer conclusiones funcionales y evolutivas sobre HLA-G de los resultados (principalmente del alelo nulo HLA-G\*01:05N), en las diferentes poblaciones del mundo.

# MATERIAL Y MÉTODOS



## **1. SELECCIÓN DE MUESTRAS.**

Para realizar este estudio, se han utilizado ciento cincuenta y cuatro individuos para los que han sido determinados sus alelos HLA-G, (dos alelos por cada individuo, uno paterno y otro materno) con cuatro dígitos de resolución. Para ciento diecisiete de ellos, también se han determinado sus alelos HLA-G no proteicos, con seis dígitos de resolución.

Entre dichos ciento cincuenta y cuatro individuos, para ciento veintidós también se han determinado sus alelos HLA-A, ciento dieciséis sus alelos HLA-A y HLA-DRB, para ochenta y tres sus alelos HLA-A, -B y -G y para ochenta y seis sus alelos HLA-A, -B, -DRB1 y -G.

## **2. OBTENCIÓN DE MUESTRAS.**

Las muestras han sido recolectadas gracias a la colaboración de la Dra. Sedeka Abd-El-Fatah-Khalil en las colectas de donación de sangre del Centro de Transfusión de la Comunidad de Madrid.

Los individuos seleccionados son donantes de sangre, sanos y no emparentados, que voluntariamente han firmado un consentimiento para participar en el estudio. Todos ellos amerindios inmigrantes en Madrid, con parámetros antropométricos, externos y visibles, típicos de su origen, posteriormente confirmado genéticamente, utilizando los genes HLA.

## **3. PREPARACIÓN DE LAS MUESTRAS.**

Una vez extraída la muestra de sangre, en tubos de 9 ml y utilizando EDTA como anticoagulante, se ha procedido al centrifugado a 4.000 rpm durante 4 minutos.

Al finalizar la centrifugación, en cada tubo se observan tres fases, hematíes la inferior, capa leucoplaquetaria, también llamada “*buffy coat*”, la intermedia y el plasma la superior.

Para obtener el ADN de cada individuo se ha separado la capa leucoplaquetaria del resto y se almacena y conserva a -20 °C hasta su extracción.

## **4. EXTRACCIÓN DE ADN GENÓMICO.**

El proceso de extracción de ADN se ha desarrollado en un extractor automático (*Nucleic Acid Extraction System, QuickGene-810* de FUJIFILM).

Tras la descongelación previa de las muestras, se ha realizado el protocolo de extracción de ADN de sangre entera gracias al kit comercial (*QuickGene Whole Blood Extraction Kit S*, FUJIFILM) realizando los siguientes pasos:

- Diluir 100 µl de sangre en 150 µl de solución salina (*PBS pH 7.4*, GIBCO) en un tubo de 1.5 ml.
- Añadir 30 µl de proteasa (*DNA whole blood Protease EDB-01*, FUJIFILM).
- Añadir 250 µl de solución de lisis (*DNA whole blood Lysis Buffer LDB-04*, FUJIFILM). Mezclar bien.
- Incubar en un baño a 56 °C hasta que desaparezcan los posibles coágulos (2-10 min).
- Añadir 250 µl de etanol absoluto frío (-20°C). Mezclar bien.
- Pasar la muestra a una columna del kit e introducirla en el extractor automático, donde será sometida a 3 ciclos de lavado (*DNA whole blood Wash Buffer WDB-03*, FUJIFILM) y posteriormente será eluída en 200 µl de una solución de elución (*DNA whole blood Elution Buffer CDB-02*, FUJIFILM).

Finalmente, para comprobar que la concentración y la pureza de la muestra han sido las adecuadas, se ha recurrido al espectrofotómetro (*Spectrophotometer ND-1000*, NANODROP), dando por válido concentraciones de entre 100-120 ng/µl y una relación A260/A280 con valor superior de 1,6 para una buena pureza.

### 5. AMPLIFICACIÓN DE SECUENCIAS HLA-G.

Posterior a la extracción del ADN de cada muestra, el proceso se ha continuado con la amplificación de la secuencia deseada, en este caso la secuencia corresponde a los exones 2, 3 y 4 del gen HLA-G.

La amplificación de una secuencia tiene como finalidad el aumento exponencial del número de copias de una secuencia determinada de ADN.

Se lleva a cabo gracias a la Reacción en Cadena de la Polimerasa (PCR)(Saiki et al., 1988), descubierta en 1986 por Kary Mullis. Esta técnica se basa en la capacidad natural de replicación del ADN a partir de una hebra molde. Para ello, se necesita además de la hebra molde, los cuatro tipos de nucleótidos (por unión de una de las cuatro bases nitrogenadas existentes; Adenina, Guanina, Timina y Citosina en el ADN), dos cebadores complementarios a la hebra que flanquean, y por tanto determinan el fragmento del ADN que será amplificado. El cebador directo se une al extremo 5' de la hebra y el inverso al extremo 3'. Y, por último, una enzima polimerasa estable y activa

a temperaturas elevadas, la más utilizada se denomina 'Taq' y se obtiene a partir del gen de la polimerasa de la bacteria *Thermus aquaticus*.

La PCR consiste en la repetición cíclica de tres etapas:

- Desnaturalización: separación de las dos hebras debido a temperaturas elevadas (95 °C aproximadamente)
- Hibridación: unión de los cebadores a sus secuencias homólogas en el ADN molde. Los cebadores son oligonucleótidos sintéticos. Para que actúen debidamente, debe haber un descenso en la temperatura y no deben ser ni autocomplementarios ni complementarios entre sí.
- Elongación: polimerización de la nueva cadena en sentido 5' → 3', gracias a la enzima ADN polimerasa (Taq polimerasa en este caso). Con temperatura de máxima actividad de 75-80 °C aproximadamente.

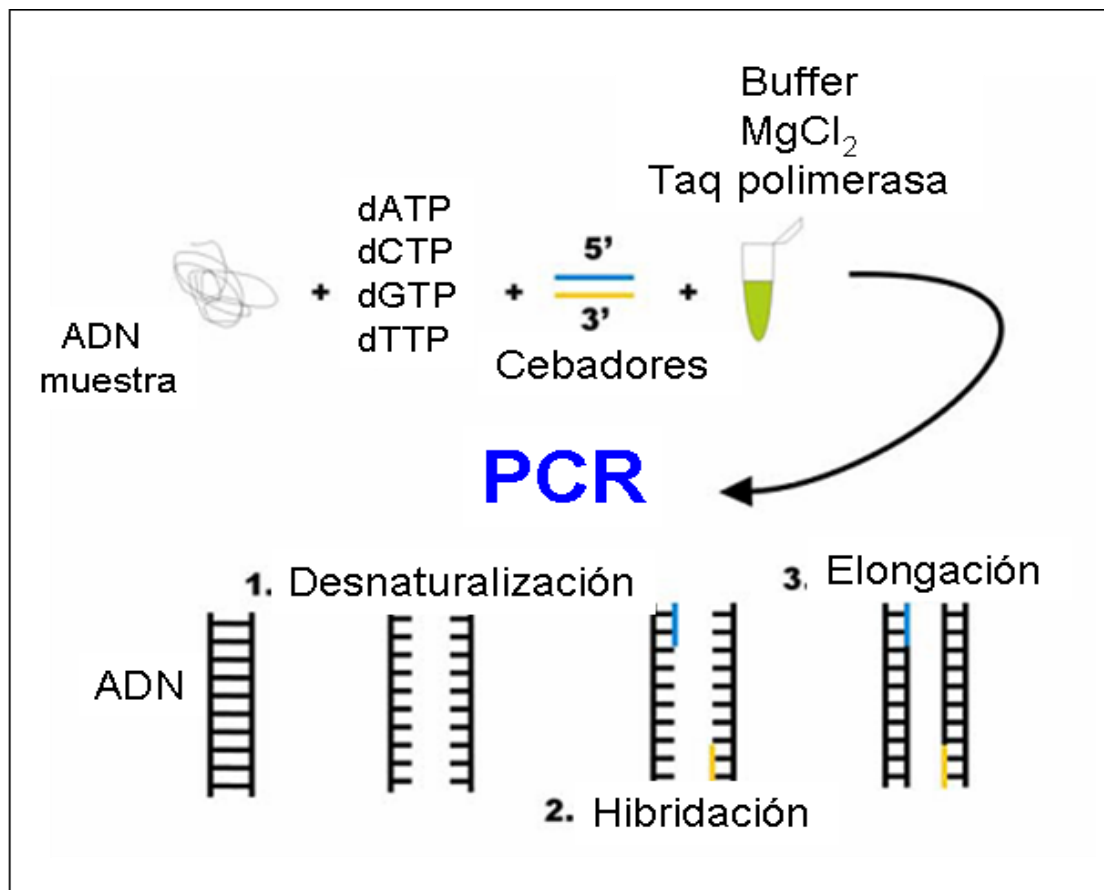


Figura M1: Esquema de la Reacción en Cadena de la Polimerasa.

Estas tres etapas ocurren dentro de un aparato llamado termociclador, capaz de elevar o disminuir la temperatura según la fase del ciclo en la que se encuentre.

La variabilidad en la región codificante del gen HLA-G, está distribuida a lo largo del exón 2, exón 3 y exón 4 (Donadi et al., 2011). Por ello se ha procedido a la amplificación de estos tres segmentos del gen con las siguientes condiciones:

EXÓN 2:

MEZCLA DE REACCIÓN			
Reactivo	[ ]	muestra (μl)	MIX (μl)
Buffer	10x	2	8
Cl <sub>2</sub> Mg	50mM	0,6	2,4
Primer 5'	10 mM	1	4
Primer 3'	10 mM	1	4
dNTPs	2.5 mM	1,6	6,4
Taq pol	5U/ml	0,25	1
H <sub>2</sub> O <sub>d</sub>	-	9,55	38,2
DNA		4	64
TOTAL:		20μl	

CONDICIONES PCR		
Nº ciclos	T (°C)	t (min)
1	95	5
35	95	0,5
	68	0,5
	72	0,67
1	72	10
-	4	(pausa)
Tiempo estimado:		73' 27"

EXÓN 3:

MEZCLA DE REACCIÓN			
Reactivo	[ ]	muestra (μl)	MIX (μl)
Buffer	10x	2	8
Cl <sub>2</sub> Mg	50mM	0,6	2,4
Primer 5'	10 mM	1	4
Primer 3'	10 mM	1	4
dNTPs	2.5 mM	1,6	6,4
Taq pol	5U/ml	0,25	1
H <sub>2</sub> O <sub>d</sub>	-	9,55	38,2
DNA		4	64
TOTAL:		20μl	

CONDICIONES PCR		
Nº ciclos	T (°C)	t (min)
1	94	2
35	94	1
	67	1,5
	72	2
1	72	10
-	4	(pausa)
Tiempo estimado:		109' 15"

EXÓN 4:

MEZCLA DE REACCIÓN			
Reactivo	[ ]	muestra (μl)	MIX (μl)
Buffer	10x	2	8
Cl <sub>2</sub> Mg	50mM	0,6	2,4
Primer 5'	10 mM	1	4
Primer 3'	10 mM	1	4
dNTPs	2.5 mM	1,6	6,4
Taq pol	5U/ml	0,25	1
H <sub>2</sub> O <sub>d</sub>	-	9,55	38,2
DNA		4	64
TOTAL:		20μl	

CONDICIONES PCR		
Nº ciclos	T (°C)	t (min)
1	95	5
35	95	0,67
	58	0,67
	72	0,67
1	72	10
-	4	(pausa)
Tiempo estimado:		76' 57"

Estas tres reacciones han tenido lugar en un termociclador *Mastercycler epgradient S* de EPPENDORF utilizando la polimerasa *Taq DNA polymerase Recombinant* 5U/ $\mu$ l de INVITROGEN, suministrada junto con un tampón (*10x PCR Rxn BufferI*, INVITROGEN) y cloruro de magnesio (*50 mM MgCl<sub>2</sub> 1ml*, INVITROGEN). Los nucleótidos: *100 mM dATP solution* (250  $\mu$ l), *100mM dTTP solution* (250  $\mu$ l), *100mM dCTP solution* (250  $\mu$ l), *100mM dGTP solution* (250  $\mu$ l) de INVITROGEN. Por último, los cebadores (suministrados por IDT) para cada exón fueron:

Exón 2:

- Directo: 5'- GAGGGTCGGGCGGGTCTCAAC -3'
- Inverso: 5'- GCATGGAGGTGGGGGTCGTGA -3'

Exón 3:

- Directo: 5'TGGGCGGGGCTGACCGAGAAGGTGG-3'
- Inverso: 5' CTCTCCTTGTGCTAGGCCAGGCTGAGA-3'

Exón 4:

- Directo: 5'-CCATGAGAGATGCAAAGTGCT-3'
- Inverso: 5'-TGCTTTCCCTAACAGACATGA-3'

### **5.1 Control de la amplificación.**

Una vez obtenido el producto de amplificación, se ha realizado un análisis electroforético para identificar y aislar el fragmento deseado.

Para ello se ha utilizado un gel de agarosa (*HR Agarose*, BIOTOOLS) al 2% en tampón TBE 1X (*UltraPure 10X TBE Buffer*, INVITROGEN). En el proceso se ha incluido un marcador de peso molecular (*DNA Molecular Weight Marker XIV - 100 base pair ladder*, ROCHE) para identificar el fragmento deseado, en torno a los 270 pares de bases (pb) cada exón, aproximadamente. Para poder observar las bandas en el gel, este se ha teñido en bromuro de etidio que emite fluorescencia al ser sometido a radiación ultravioleta.



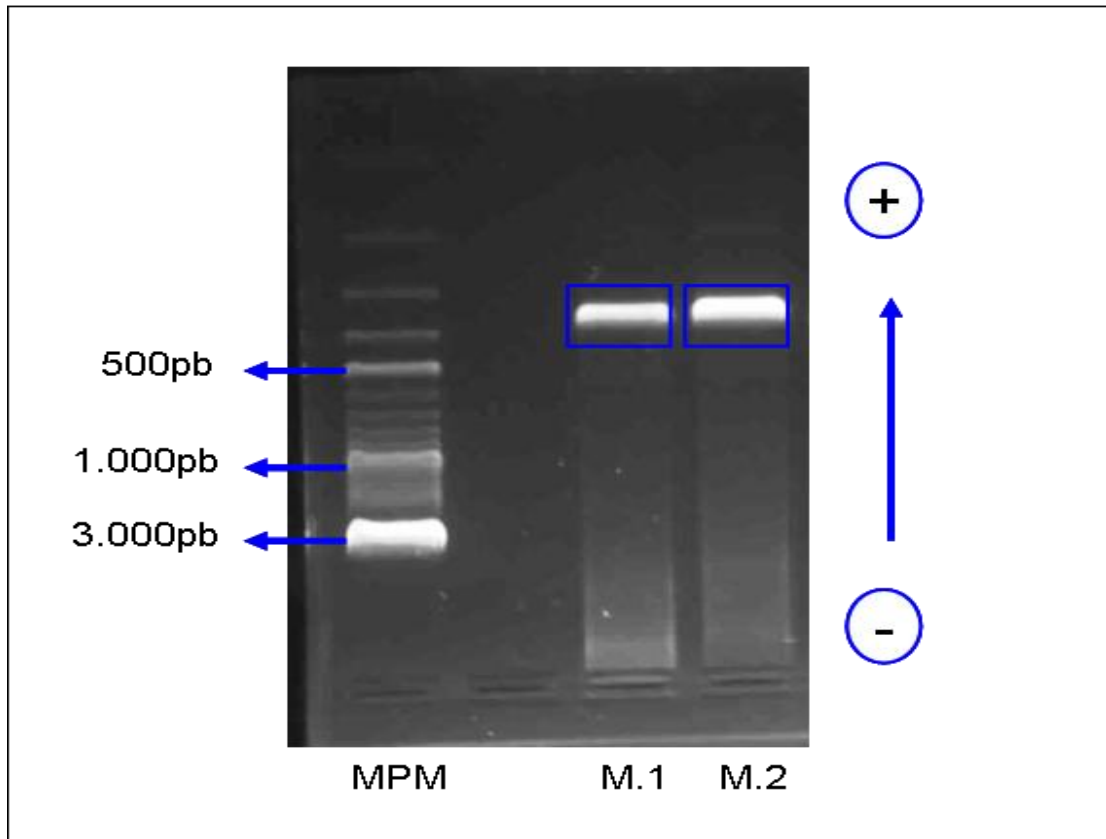


Figura M2: Imagen gel de agarosa. Figuran las bandas de ADN del exón 2 de HLA-G para la muestra 1 (M.1) y muestra 2 (M.2). En el primer carril se observa el marcador de peso molecular (MPM).

## 5.2 Purificación de las bandas.

Las bandas se recortan y purifican con un kit comercial *MiniElute Gel Extraction Kit 250*, QIAGEN. Se sigue el siguiente protocolo:

- Añadir 300 µl de *Buffer QG Solubilization Buffer250 ml* (QIAGEN) en el tubo que contiene la banda de gel recortada e introducir en un baño a 56°C hasta que se disuelva el gel (5-10 minutos).
- Pasar la muestra ya disuelta a una columna del kit.
- Centrifugar 1 minuto a 14.000 rpm. Decantar el filtrado.
- Añadir 500 µl de *Buffer QG*.
- Centrifugar 1 minuto a 14.000 rpm. Decantar el filtrado.
- Añadir 750 µl de *Buffer PE - Wash Buffer 275 ml* (QIAGEN) y dejar actuar 5 minutos.
- Centrifugar 1 minuto a 14.000 rpm. Decantar el filtrado.
- Repetir el paso anterior.
- Colocar la columna en un tubo de 1.5 ml.
- Añadir 50 µl de agua destilada.

- Centrifugar 1 minuto a 14.000 rpm.
- Tirar la columna y almacenar el ADN purificado a 4°C o a -20°C hasta su uso para la secuenciación directa.

## **6. SECUENCIACIÓN DE LOS FRAGMENTOS AMPLIFICADOS.**

Las reacciones de secuenciación están basadas en el método de Sanger (Sanger et al., 1977), que utiliza dos técnicas ya descritas, la polimerización de ADN a partir de un molde a través de la PCR y la técnica de electroforesis.

Esta PCR tiene la peculiaridad de, además de incorporar los nucleótidos normales, utiliza dideoxinucleótidos, que carecen del grupo hidroxilo del carbono 3' (necesario para incorporar nuevos nucleótidos durante la polimerización). Por ello, cada vez que se incorporan a la cadena de forma aleatoria, se interrumpe la polimerización. De tal forma que, al finalizar el proceso, habrá fragmentos de todas las longitudes posibles y todos ellos tendrán un dideoxinucleótido en su extremo 3'. Otra característica de estos dideoxinucleótidos es que están marcados con un fluorocromo, cada uno de un color distinto. La reacción de secuenciación se desarrolla en cuatro tubos independientes, en cada uno de los cuales se añade un tipo de dideoxinucleótidos marcado. Los productos resultantes, se someten a electroforesis en cuatro calles paralelas (una por dideoxinucleótido).

Una vez revelado el gel se podrá leer la secuencia siguiendo un orden ascendente de tamaño de los fragmentos distribuidos a lo largo de las cuatro calles del gel.

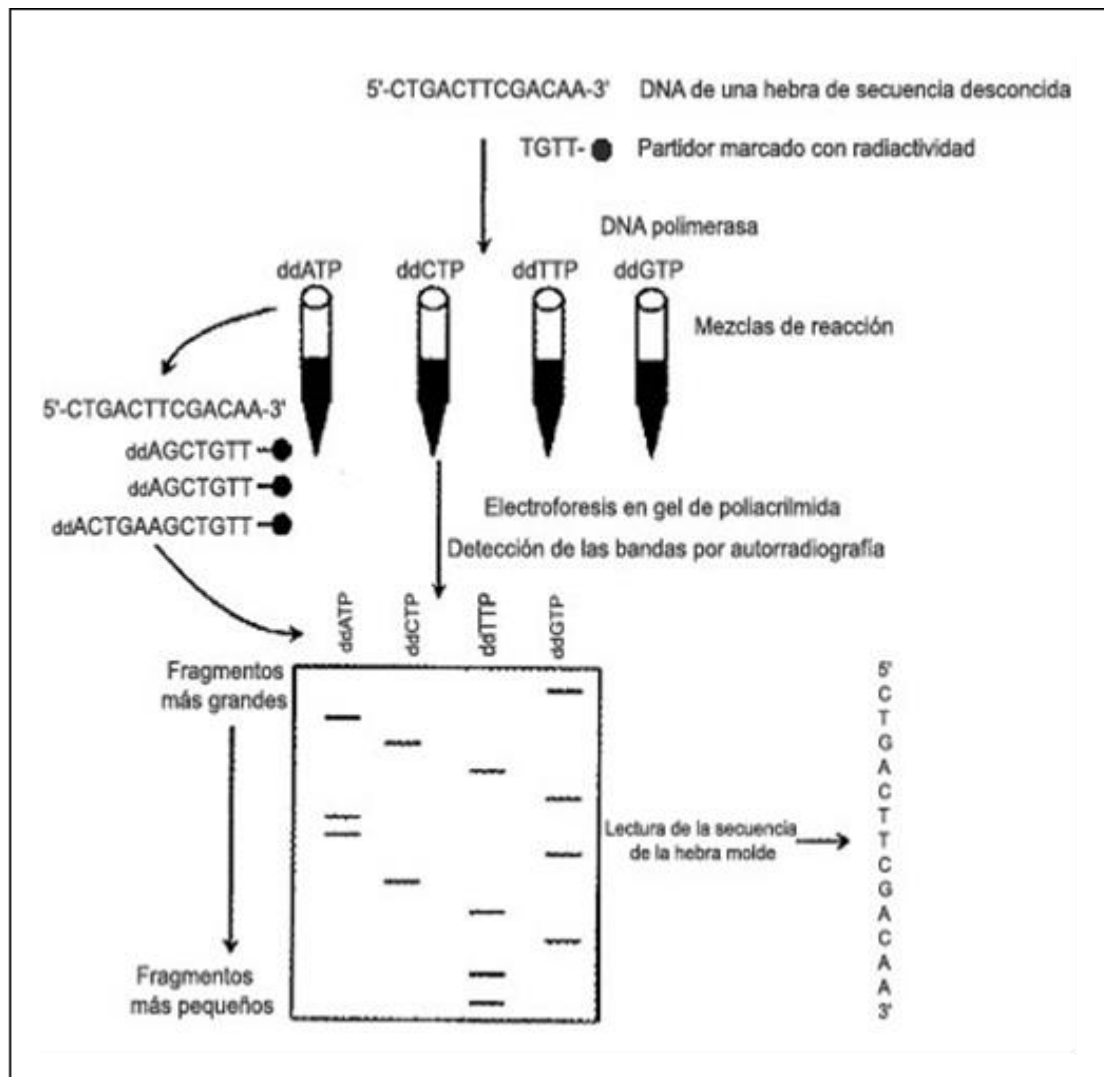


Figura M3: Esquema del proceso seguido para el método de Sanger.

Actualmente, las reacciones de secuenciación se llevan a cabo en secuenciadores automáticos de ADN, en los que las reacciones se desarrollan en un mismo tubo. El marcaje de cada dideoxinucleótido es detectado directamente por el secuenciador.

Para llevar a cabo este proceso de secuenciación, se enviaron las muestras a la empresa SECUGEN S.L. ([www.secugen.es](http://www.secugen.es)), situada en el Centro de Investigaciones Biológicas (CIB-CSIC) de Madrid.

Antes de ser enviadas, se prepararon las muestras utilizando 15 µl del ADN amplificado y purificado con una concentración de alrededor de 10 ng/ml, y 1.5 µl del mismo cebador que se usó en la amplificación, pero con una concentración de 5 mM.

Una vez que se ha realizado este proceso en SECUGEN ([www.secugen.es](http://www.secugen.es)), nos envían las secuencias.

## 7. DETERMINACIÓN DE ALELOS HLA-G.

El gen HLA-G es poco polimórfico (Jeong et al., 2014; Lynge et al., 2014), y en la actualidad se conocen 53 alelos para este gen y 18 proteínas (Gonzalez-Galarza et al., 2011). Dado que es un número bajo respecto al resto de genes de clase I, la determinación de los alelos se ha realizado de forma manual.

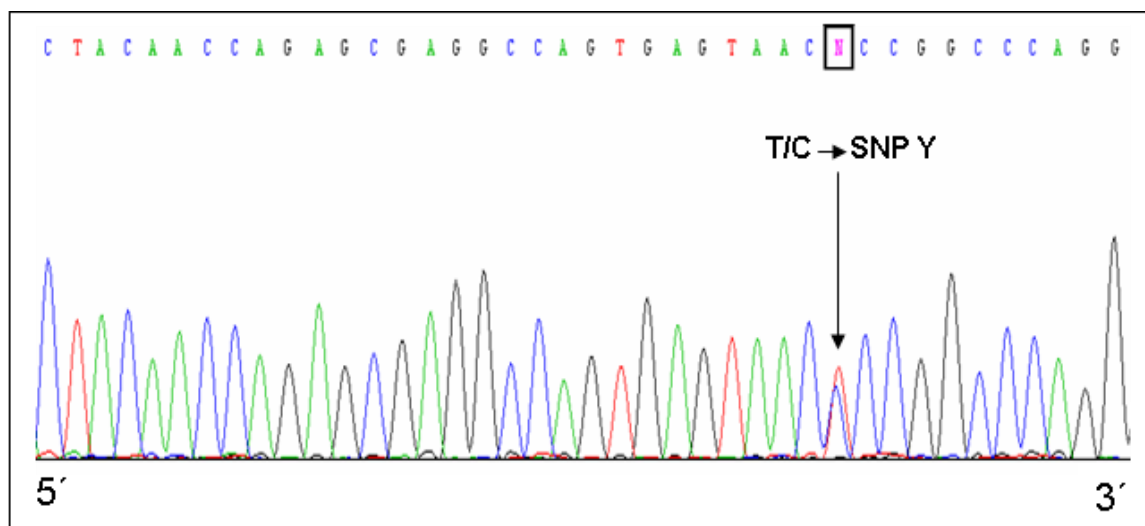


Figura M4: *Electroferograma de una secuencia de HLA-G. Se observa un doble pico o ambigüedad entre una T y una C que corresponde a un SNP denominado con la letra Y.*

Como se observa en la figura M4, las secuencias recibidas pueden tener ambigüedades que han de ser esclarecidas para poder determinar que par de alelos tiene cada individuo, dado que estas ambigüedades indican que el individuo es heterocigoto para este gen.

Por cada individuo se ha obtenido dos secuencias, la 5' y la 3', para cada exón. Ambas se han alineado gracias al programa MEGA 7 (Kumar et al., 2015). Una vez alineadas, a partir de estas, se crea una secuencia consenso con los extremos del exón limpios y con los SNPs localizados, si es que los hubiera, nombrando cada uno de ellos con su letra correspondiente (ver tabla de correspondencia de ambigüedades, Anexo II)

Al terminar este proceso se obtiene, por cada individuo, tres secuencias consenso (una por cada exón) que son unidas en una secuencia consenso única. Esta secuencia se ha comparado, alineándola, con cada una de las secuencias de los alelos ya descritos (ver tabla alelos HLA-G, Anexo IV).

En el caso de un individuo homocigoto (ausencia de ambigüedades), la secuencia consenso solo coinciden por completo con un único alelo (el mismo heredado por vía

paterna y materna) que es el alelo que lleve en el locus HLA-G en ambos cromosomas homólogos.

En el caso de un individuo heterocigoto (presencia de ambigüedades), la secuencia consenso coincide con dos alelos distintos. Cada uno de esos dos alelos lleva una base nitrogenada en la posición del SNP determinado, es decir, habiéndose detectado un SNP “M”, un alelo llevará una Adenina (A) y el otro alelo una Citosina (C) en la misma posición. Por lo tanto, se compara la secuencia consenso con la lista de alelos posibles hasta que se identifiquen el par de alelos complementarios entre sí.

Todo esto se ha realizado dentro del mismo programa MEGA 7 (Kumar et al., 2015).

## **8. ANÁLISIS ESTADÍSTICOS.**

Una vez obtenidos los alelos de nuestra población de estudio, se ha procedido a su análisis estadístico. Para ello se ha contado con determinados programas informáticos comentados a continuación.

### **8.1 Cálculo de frecuencias alélicas.**

La frecuencia de cada alelo se ha obtenido mediante el programa Arlequín v2.0, que realiza un recuento de cada variante alélica (alelo) y lo divide entre el número total de variantes alélicas encontradas en nuestra población de estudio (Schneider et al., 2000).

### **8.2 Estimación de los haplotipos extendidos y desequilibrio de ligamiento.**

El cálculo de haplotipos, su desequilibrio de ligamiento, la frecuencia de los mismos y el valor estadístico de significación  $p$  para cada asociación, se ha realizado también mediante el programa informático Arlequín v2.0 (Schneider et al., 2000) y Arlequín v3.0 (Excoffier et al., 2005)(para el desequilibrio de ligamiento entre distintos alelos). Este programa, combina el algoritmo de expectación-maximización (EM') con el método de máxima probabilidad (*maximum-likelihood*). Este es un proceso iterativo, que persigue estimados de máxima probabilidad de frecuencias haplotípicas, obtenidos a partir de datos genotípicos de varios loci cuando su distribución gamética es desconocida. El principio del algoritmo EM' es el siguiente:

- Cálculo inicial aleatorio de frecuencias haplotípicas.
- Uso de estos estimados para computar las frecuencias genotípicas esperadas para cada

fenotipo, asumiendo el equilibrio de Hardy-Weinberg (paso de expectación).

- Las frecuencias genotípicas relativas, sirven como pesos para los dos haplotipos que se constituyen en un procedimiento de conteo génico, dando lugar a nuevos estimados de frecuencias haplotípicas (paso de maximización).
- Se repiten los pasos 2 y 3 hasta que las frecuencias de los haplotipos alcancen el equilibrio.

Dos o más alelos de loci diferentes, se consideran asociados o en desequilibrio de ligamiento, cuando aparecen juntos en un mismo cromosoma con mayor frecuencia de la que cabría esperar del producto de las frecuencias génicas de los alelos implicados. El desequilibrio de ligamiento (D) se define matemáticamente como:

$$D_{ab} = F_{Hab} - (F_a \times F_b)$$

donde  $F_{Hab}$  es la frecuencia del haplotipo  $ab$  en la población,  $F_a$  es la frecuencia alélica de  $a$ , y  $F_b$  es la frecuencia alélica de  $b$ .

Los haplotipos con  $D < 0$  tienden a no estar juntos formando otros haplotipos, mientras que las asociaciones con  $D > 0$  indican que los alelos que componen los haplotipos tienden a formar estos y no otros. Los haplotipos con valores de  $D > 0$  pueden, a su vez, tener una probabilidad estadísticamente significativa para un grado de libertad de  $p < 0,05$ , es decir, existe un 95% de seguridad de que la asociación de alelos que forman el haplotipo es real y no se debe al azar. En este trabajo se describen aquellos haplotipos de dos loci con desequilibrio de ligamiento positivo ( $D > 0$ ) y  $p < 0,05$ .

Por otro lado, también se ha usado el desequilibrio de ligamiento normalizado ( $D'$ ), que ajusta la influencia de las frecuencias alélicas variables.  $D' = D/D_{max}$ , y  $D_{max}$  = máximo  $D$  dadas las frecuencias alélicas. Fluctúa entre  $-1$  y  $+1$ . Es útil en la comparación de diferentes valores de  $D$ .

Para el cálculo de los haplotipos extendidos que definen una población, se utiliza el algoritmo empleado por el programa Arlequín v2.0 (Schneider et al., 2000) de EM'.

### **8.3 Test de Fisher.**

Para comparar la frecuencia de los alelos nulos (HLA-G\*01:05N) de la población de estudio con otras poblaciones mundiales, se ha usado el Test de Fisher. La significación estadística ha sido definida como  $p < 0,05$ . Este análisis ha sido ejecutado gracias al

programa estadístico SPSS/PC (versión 17.0 para Windows; SPSS, Inc., Chicago, IL, USA)

## **9. CARACTERIZACIÓN DE LOS NUEVOS ALELOS.**

En el proceso de determinación de los alelos HLA-G en la población amerindia, se han encontrado dos alelos nuevos, por lo que, no habían sido descritos con anterioridad.

Todo el protocolo de actuación seguido con estos alelos nuevos, ha sido el mismo que se ha mencionado anteriormente hasta la determinación de los alelos HLA-G, donde se ha observado que las secuencias consenso de los dos alelos encontrados en dos individuos distintos, no se alineaban correctamente con ninguna de las secuencias ya descritas para el locus HLA-G (53 alelos hasta la fecha y 18 proteínas) (Robinson et al., 2015) (ver tabla I5).

Los dos individuos, a los que pertenecían ambas muestras de ADN, eran de nacionalidades distintas. Así, el primer individuo era un hombre procedente de Ecuador y el segundo una mujer procedente de República Dominicana.

### **9.1 Repetición de la secuenciación directa.**

Este paso se ha realizado para descartar posibles errores cometidos a lo largo de todo el proceso, tanto errores técnicos como humanos.

Para ello, se volvieron a realizar tres PCR para el individuo uno (proveniente de Ecuador) una para cada exón, y dos PCR para el individuo dos (proveniente de República Dominicana) para los exones 2 y 3, dado que no se había encontrado ningún cambio en la secuencia del exón 4. Seguidas de la electroforesis y la purificación de las bandas, se enviaron las muestras obtenidas de nuevo a SECUGEN ([www.secugen.es](http://www.secugen.es)).

### **9.2 Repetición de la determinación de los alelos.**

Una vez recibidas las secuencias, se ha procedido a la alineación, de nuevo, con cada uno de las secuencias de los alelos ya descritos. Viendo que el resultado era el mismo al obtenido la vez primera, se ha iniciado un nuevo procedimiento de verificación.

### **9.3 Secuenciación mediante clonaje del producto amplificado.**

Cuando los productos de amplificación son directamente secuenciados, las secuencias de ADN resultantes pueden presentar ambigüedades, como ya se ha visto, por lo tanto, para obtener la secuencia completa de los nuevos alelos ha sido necesario someter a los productos de amplificación a un proceso de clonación antes de ser definitivamente secuenciados.

La clonación se basa en la capacidad natural que tienen algunas bacterias para incorporar pequeñas moléculas circulares de ADN (plásmidos) presentes en su entorno, en un proceso denominado ‘transformación bacteriana’ (Sambrook et al., 1989). Este procedimiento permite obtener una secuencia única, útil para aquellos casos en que los alelos no pueden separarse por grupos. Los plásmidos contienen genes que aportan funciones de mejora para la supervivencia de la bacteria, como por ejemplo la resistencia a determinados antibióticos, y son replicados en el interior de la bacteria transmitiéndose, por tanto, a sus descendientes. En general, la eficacia de la transformación es bastante reducida, y la mayoría de las bacterias que la realizan con éxito incorporan un único plásmido. Estas propiedades bacterianas, sumadas a la posibilidad de diseñar plásmidos artificiales (denominados vectores de forma genérica) e insertar en ellos fragmentos de ADN mediante técnicas de ingeniería genética, hacen de la clonación una herramienta muy útil para aislar y caracterizar secuencias de ADN de forma individual.

Para llevar a cabo la clonación, se ha utilizado el sistema de *pGEM-T Easy* (PROMEGA). La utilidad de este vector se debe a la presencia de un gen de resistencia a ampicilina, que permite seleccionar las bacterias que han incorporado el plásmido, con o sin fragmento. Por otro lado, el lugar de inserción se encuentra en mitad del gen *lacZ*. La integridad de dicho gen, cuya proteína metaboliza el X-gal, les confiere un color azulado. Esto permite seleccionar las bacterias que han incorporado el plásmido con el fragmento insertado (serán de color blanco, dado que, el gen *lacZ* se ha visto interrumpido por el inserto). Además, el plásmido está linearizado y presenta extremos monocatenarios en 5’ formados por un nucleótido de timina, de manera que se facilita la inserción de cualquier fragmento del tamaño apropiado que posea, a su vez, extremos monocatenarios constituidos por un nucleótido de adenina en el extremo 3’.



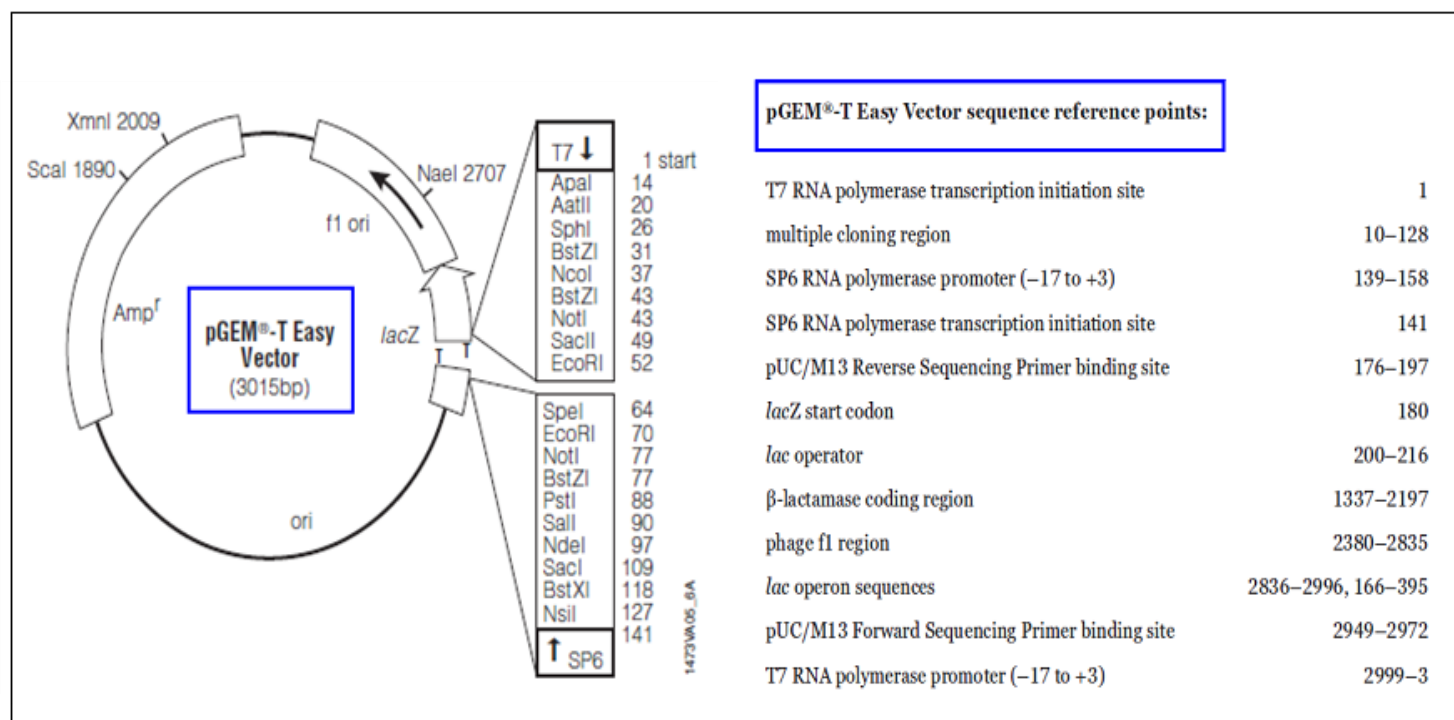


Figura M5: Esquema del vector utilizado en el proceso de clonación.

El procedimiento seguido consiste en:

- Preparar la reacción de ligado mezclando 5 µl de tampón (*2x Rapid Ligation Buffer*, PROMEGA), 1 µl de plásmido (*pGEM-T Easy Vector*, PROMEGA), 3 µl de producto de amplificación y 1 µl de ligasa (*T4 DNA Ligase*, PROMEGA).
- Incubar 90 minutos a temperatura ambiente o bien durante 12-16 horas a 4°C.
- Añadir 50 µl de células competentes (*JM109 Competent Cells, High efficiency*, PROMEGA) a un tubo conteniendo 2 µl de la mezcla de ligado.
- Incubar en hielo durante 20 minutos.
- Incubar en un baño a 42°C durante 40-45 segundos.
- Incubar en hielo durante 2 minutos.
- Añadir 950 µl de Medio SOC, que contiene 30.7 gr/l de Medio SOB (*SOB MEDIA*, PRONADISA) y Glucosa 20 mM (*DEXTROSA*, PRONADISA).
- Incubar a 37°C en agitación (150 rpm) durante 90 minutos.
- Sembrar 100 µl en placas con Medio LB Agar conteniendo ampicilina, IPTG y XGAL (2 placas por muestra). 1 litro de medio contiene 35 gr de LB Agar (*LB AGAR LENNOX*, PRONADISA), 10 ml de ampicilina 10 mg/ml (*Ampicillin Sodium SALT - Irradiated*, GIBCO), 0.5 ml de IPTG 1M (*IPTG Dioxane-Free*, PROMEGA) y 1.6 ml de X-Gal 50 mg/ml (*X-GAL*, PROMEGA).

- Incubar a 37°C durante 16-24 horas.
- Opcionalmente, las placas pueden mantenerse a 4°C durante unas horas después de la incubación para favorecer la aparición de color azul en las colonias “negativas”.

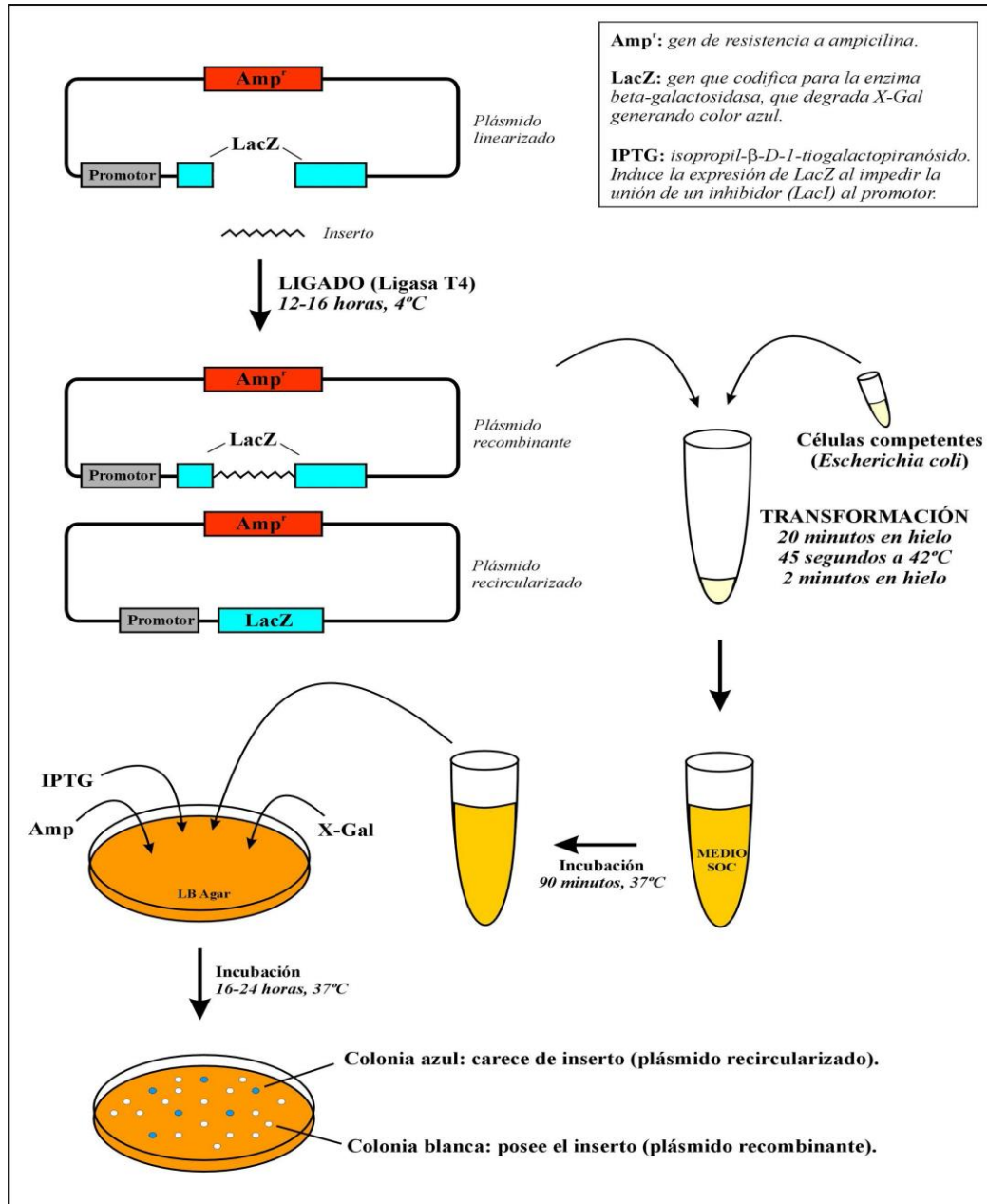


Figura M6: Esquema del proceso de clonación.

#### 9.4 Obtención de las secuencias de ADN de los fragmentos clonados.

Una vez incubadas las placas, se seleccionan diez colonias blancas de cada una (20 colonias por muestra) para extraer el ADN bacteriano mediante el siguiente procedimiento:

- Resuspender la colonia en 20 µl de agua destilada en un tubo de 0.2 ml.
- Incubar durante 5 minutos a 95°C en un termociclador.
- Incubar durante 5 minutos en hielo.
- Centrifugar 3 minutos a 12000 rpm.
- Pasar el sobrenadante (aproximadamente 15 µl) a un tubo de 1.5 ml. Conservar a -20°C.

Posteriormente, se han amplificado las secuencias clonadas utilizando como molde el ADN extraído de las bacterias (que incluye el cromosoma bacteriano y el plásmido con el inserto). La PCR se ha desarrollado en las mismas condiciones en las que se ha amplificado el ADN genómico. Los productos de amplificación han sido sometidos a electroforesis para separar el fragmento correspondiente, que finalmente ha sido secuenciado por SECUGEN ([www.secugen.es](http://www.secugen.es)), siguiendo el mismo procedimiento explicado en el apartado 6. Finalmente, se han obtenido las secuencias exactas y sin ambigüedades, de los nuevos alelos.

### **9.5 Deposito de los nuevos alelos de HLA-G.**

Al descubrirse estos nuevos alelos, (ver Resultados, apartado 5) se ha utilizado el programa Sequin 13.70 (<http://www.ncbi.nlm.nih.gov/Sequin/>) para recibir numeración identificativa única del ADN de los alelos descubiertos nuevos.

### **9.6 Determinación de las distancias genéticas.**

Una vez determinada la secuencia nucleotídica de ambos alelos, se han calculado las distancias genéticas entre los dos alelos nuevos y el resto de alelos determinados hasta la fecha.

La distancia genética entre dos secuencias de ADN se calcula en base a las diferencias que se observan entre ellas. Si se compararan dos secuencias de 1.000 nucleótidos que se diferencian en 6 de ellos, la distancia p sería 0,006. Según esto, la distancia entre alelos que tienen la misma secuencia de nucleótidos es cero.

En este caso, el cálculo se ha realizado gracias al programa MEGA 7 (Kumar et al., 2015).

### **9.7 Determinación de las secuencias aminoacídicas de los alelos nuevos.**

Finalmente, se ha procedido a determinar la secuencia aminoacídica, gracias de nuevo, al programa MEGA 7 (Kumar et al., 2015). Aportando la secuencia nucleotídica, el programa realiza la conversión de ADN a proteína.

# RESULTADOS



# 1. DETERMINACIÓN DE LOS ALELOS HLA-G EN LA POBLACIÓN AMERINDIA INMIGRANTE EN MADRID.

Se han obtenido los alelos HLA-G, por técnicas de genética molecular, con cuatro dígitos de resolución, en una muestra poblacional de ciento cincuenta y cuatro individuos y con seis dígitos de resolución (alelos no proteicos), en una muestra de ciento diecisiete individuos. En ambos casos son individuos sanos no relacionados. Éstos fueron donantes de sangre, voluntarios para participar en el estudio. Todos ellos amerindios inmigrantes en Madrid, con parámetros antropométricos, externos y visibles típicos de su origen, y confirmado genéticamente, utilizando los genes HLA. (Ver lista de alelos en Anexos VI)

La muestra tomada es muy heterogénea, los individuos pertenecen a dieciséis países amerindios distintos.

**Tabla R1:** Países de origen de los individuos de la muestra, con sus porcentajes.

País de origen	n	%
Argentina	7	4,5
Uruguay	4	2,6
Ecuador	45	28,8
Bolivia	10	6,4
Perú	21	13,5
Cuba	1	0,6
Paraguay	2	1,3
Guatemala	1	0,6
Venezuela	2	1,3
Colombia	42	26,9
Chile	4	2,6
Brasil	5	3,2
Méjico	4	2,6
R.Dominicana	6	3,8
Nicaragua	1	0,6
Honduras	1	0,6

*n; número de individuos.*



Figura R1: Mapa de Centro y Suramérica con el número de individuos de cada país.

En la tabla R1 se observa como el país más representado en la muestra de estudio es Ecuador, con un porcentaje del 28,8%, seguido de Colombia (26,9%), Perú (13,5%) y Bolivia (6,4%). Los menos representados, con un solo individuo, son Guatemala, Cuba, Nicaragua, Venezuela y Honduras.

En la figura R1 se muestra de una forma más gráfica los países que intervienen en el estudio con el número de individuos de cada uno.



## **2. CÁLCULO DE LAS FRECUENCIAS ALÉLICAS HLA-G EN LA POBLACIÓN AMERINDIA INMIGRANTE EN MADRID.**

Las frecuencias alélicas del locus HLA-G se han obtenido gracias al programa informático Arlequín v2.0. En primer lugar, se han obtenido las frecuencias con cuatro dígitos de resolución y posteriormente con seis dígitos de resolución, dando como resultado las siguientes dos tablas.

**Tabla R2:** Frecuencias alélicas HLA-G, con cuatro dígitos de resolución, en amerindios inmigrantes en Madrid. (n=154). En negrita se marca la presencia del alelo nulo.

Alelos	Frecuencias %
HLA-G*01:01	71,8
HLA-G*01:04	18,5
HLA-G*01:03	5,8
HLA-G*01:06	2,9
<b>HLA-G*01:05N</b>	<b>1,0</b>

*n; número de individuos.*

**Tabla R3:** Frecuencias alélicas HLA-G, con seis dígitos de resolución, en amerindios inmigrantes en Madrid. (n=117). En negrita se marca la presencia del alelo nulo.

Alelos	Frecuencias %	Alelos	Frecuencias %
HLA-G*01:01:01	31,2	HLA-G*01:01:19	0,4
HLA-G*01:01:02	23,1	HLA-G*01:01:20	1,3
HLA-G*01:01:03	4,3	HLA-G*01:03:01	7,7
HLA-G*01:01:05	0,9	HLA-G*01:04:01	12,8
HLA-G*01:01:06	2,1	HLA-G*01:04:03	1,7
HLA-G*01:01:07	0,4	HLA-G*01:04:04	3,8
HLA-G*01:01:08	3,0	HLA-G*01:04:05	0,4
HLA-G*01:01:15	0,4	<b>HLA-G*01:05N</b>	<b>1,3</b>
HLA-G*01:01:17	0,9	HLA-G*01:06	0,4

*n; número de individuos.*

Teniendo en cuenta los alelos HLA-G no proteicos, con seis dígitos de resolución, de los 53 descritos en la actualidad (Robinson et al., 2015), solo dieciocho están representados en esta población. Por otro lado, con cuatro dígitos de resolución, solo se han encontrado cinco alelos diferentes. Dentro del grupo de alelos HLA-G\*01:01, se han detectado los once alelos siguientes (más de la mitad del total de alelos encontrados) HLA-G\*01:01:01, HLA-G\*01:01:02, HLA-G\*01:01:03, HLA-

G\*01:01:05, HLA-G\*01:01:06, HLA-G\*01:01:07, HLA-G\*01:01:08, HLA-G\*01:01:15, HLA-G\*01:01:17, HLA-G\*01:01:19, HLA-G\*01:01:20

El primer alelo descrito por Geraghty, el HLA-G\*01:01 (salvaje) (Geraghty et al., 1987), es predominante en casi todas las poblaciones, con una frecuencia variable de un 83,0% en africanos de Ghana a un 3,9% en iraníes (Park et al., 2012). De tal manera, en ambas tablas se observa como el alelo no proteico dominante en la población amerindia inmigrante en Madrid es el HLA-G\*01:01:01 (31,2%), con seis dígitos de resolución, y HLA-G\*01:01 (71,8%), con cuatro dígitos, al igual que ocurre en otras poblaciones amerindias como los mayas y uros (Arnaiz-Villena et al., 2013)(55) y en brasileños, aunque es una población con mucho más mestizaje (Castelli et al., 2007). Seguidamente, el alelo no proteico más frecuente, con seis dígitos de resolución, es el HLA-G\*01:01:02 (23,1%) y con cuatro dígitos de resolución, se cambia a otro grupo de alelos, HLA-G\*01:04 (18,5%). Se tiene en cuenta que, con seis dígitos de resolución, el tercero más frecuente ya es de este mismo grupo, HLA-G\*01:04:01 (12,8%).

En definitiva, los grupos más frecuentes en esta población son HLA-G\*01:01 y HLA-G\*01:04, como ocurre en muchas otras poblaciones mundiales como coreanos (Park et al., 2012) o en africanos de Shona (Matte et al., 2000).

La frecuencia del alelo nulo HLA-G\*10:05N está comentada más adelante. (ver apartado 4)

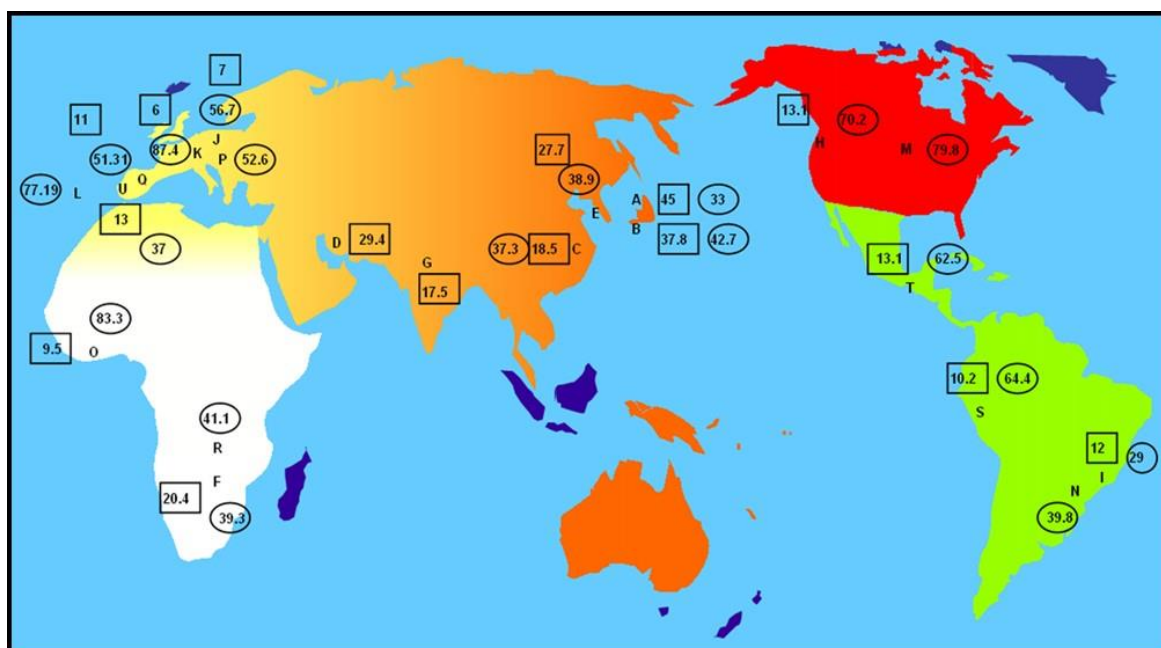


Figura R2: Patrón de las frecuencias más altas en el mundo de los alelos HLA-G\*01:01 y HLA-G\*01:04. Las frecuencias del alelo HLA-G\*01:01 están marcadas con un círculo, mientras que

*las frecuencias del alelo HLA-G\*01:04, con un recuadro. Las frecuencias alélicas de cada población se especifican a continuación:*

- A- Japon pop 14: G\*01:04→ 45%, G\*01:01→ 33% (Yamashita T et al. 1996)
- B- Japon pop 15: G\*01:04→ 37.8%, G\*01:01→ 42.7% (Ishitani et al. 1999c)
- C- China Zhejiang Provincia Han: G\*01:04→18.5%, G\*01:01→ 37.3% (Yan et al. 2006)
- D- Iran Tehran pop 2: G\*01:04→ 29.4% (Rahimi et al. 2010)
- E- Corea del Sur pop 7: G\*01:04→ 27.7%, G\*01:01→38.9% (Park et al. 2007)
- F- Zimbabue Harare Shona pop 2: G\*01:04→20.4%, G\*01:01→ 39.3% (Matte et al. 2000)
- G- India Lucknow pop 2: G\*01:04→ 17.5% (Abbas et al. 2004)
- H- USA Americanos africanos pop 6: G\*01:04→ 13.1%, G\*01:01→ 70.2% (Ishitani et al. 1999b)
- I- Brasil Parana mezcla: G\*01:04→ 12%, G\*01:01→ 29% (Pirri et al. 2009)
- J- Dinamarca pop2: G\*01:04→ 7% (Hviid et al. 1997), G\*01:01→ 56.7% (Steffensen et al. 1998)
- K- Alemania pop5: G\*01:04→ 6%, G\*01:01→ 87.4% (van der Ven et al. 1998)
- L- Portugal Azores Isla Terceira: G\*01:01→ 77.19% (Gonzalez-Galarza et al. 2015)
- M- USA Hutterite Dakota del Sur: G\*01:01→ 79.8% (Ober et al. 1996)
- N- Brasil Sao Paulo pop 2: G\*01:01→ 39.8% (Castelli et al. 2007)
- O- Ghana: G\*01:04→ 9.5%, G\*01:01→ 83.3% (Ishitani et al. 1999a)
- P- Polonia pop 2: G\*01:01→ 52.6% (Sipak-Szmigiel et al. 2009)
- Q- España pop 2: G\*01 :04→ 11%, G\*01:01→ 51.31% (Suarez et al. 1997)
- R- Zambia: G\*01:01→ 41.1% (Shao et al. 2006)
- S- Lago Titikaka Uros: G\*01:04 → 10.2%, G\*01:01→ 64.4% (Arnaiz-Villena et al. 2013)
- T- Guatemala Mayas: G\*01:04→ 13.1%, G\*01:01→ 62.5% (Arnaiz-Villena et al. 2013)
- U- Portugal: G\*01:04→13%, G\*01:01→ 37% (Alvarez et al. 1999)

Las frecuencias más altas del alelo HLA-G\* 01:01se encuentran en las poblaciones de Alemania, Ghana y Hutterite en Dakota del Sur, con una frecuencia de 87,4%, 83,3% y 79,8%, respectivamente.

En cuanto al alelo HLA-G\*01:04, las frecuencias más altas se encuentran en la población japonesa (A y B), iraní y surcoreana, con una frecuencia del 45% y 37,8%, 29,4% y 27,7%, respectivamente (figura R2). Es importante mencionar que en Europa no se ha detectado ninguna frecuencia elevada para este alelo.

Las poblaciones amerindias de Brasil, Uros y Mayas, presentan frecuencias más elevadas para el alelo HLA-G\*01:01 (29%, 64,4% y 62,5% respectivamente) que para el alelo HLA-G\*01:04 (no determinado, 10,2% y 13,1% respectivamente) como ocurre en el caso de amerindios inmigrantes en Madrid.

### **3. ANÁLISIS DE LOS HAPLOTIPOS EXTENDIDOS EN LA POBLACIÓN AMERINDIA INMIGRANTE EN MADRID.**

#### **3.1 Loci HLA-A y -G.**

Las frecuencias del haplotipo biloci HLA-A-G, se han calculado gracias al programa Arlequín v2.0 (Schneider et al., 2000). En la tabla R4 se muestran los trece haplotipos más frecuentes en la población de estudio, más las frecuencias de los haplotipos que incluyen el alelo nulo HLA-G\*01:05N.

**Tabla R4:** Frecuencias haplotípicas de los loci HLA-A y HLA-G, con cuatro dígitos de resolución, en amerindios inmigrantes en Madrid. (n=122). En negrita se marca marcan los haplotipos que porten el alelo nulo.

<b>Loci HLA-A y -G</b>	<b>Frecuencias %</b>	<b>Loci HLA-A y -G</b>	<b>Frecuencias %</b>
A*02:01 G*01:01	24,6	A*01:01 G*01:06	3,3
A*24:02 G*01:04	13,2	A*68:02 G*01:01	2,9
A*03:01 G*01:01	6,6	A*32:01 G*01:01	2,9
A*68:01 G*01:01	5,2	A*30:02 G*01:01	2,4
A*29:02 G*01:01	5,0	A*26:01 G*01:01	2,4
A*23:01 G*01:04	3,3	<b>A*30:01 G*01:05N</b>	<b>0,8</b>
A*11:01 G*01:01	3,3	<b>A*31:01 G*01:05N</b>	<b>0,4</b>
A*02:11 G*01:01	3,3		

*n*; número de individuos.

El cálculo del valor de desequilibrio de ligamiento entre estos dos loci, se ha realizado gracias al programa Arlequín v3.0 (Excoffier et al., 2005). Diez haplotipos se han encontrado cuyos alelos están en desequilibrio de ligamiento positivo ( $D > 0$ ) y  $p < 0.05$  (Tabla R5).

**Tabla R5:** Desequilibrio de ligamiento entre los alelos de los haplotipos biloci HLA-A y HLA-G, con cuatro dígitos de resolución, en amerindios inmigrantes en Madrid. (n=122). En negrita se marcan los valores de desequilibrio de ligamiento positivos que porten el alelo nulo.

Loci HLA-A y -G	D	D'
A*02:01 G*01:01	0,07	0,94
A*03:01 G*01:01	0,02	1,00
A*29:02 G*01:01	0,01	1,00
A*23:01 G*01:04	0,03	1,00
A*24:04 G*01:04	0,00	1,00
A*24:03 G*01:04	0,00	1,00
A*24:02 G*01:04	0,10	1,00
A*24:25 G*01:04	0,00	1,00
A*31:01 G*01:03	0,02	0,60
A*23:03 G*01:03	0,00	1,00
A*31:12 G*01:03	0,00	1,00
A*02:05 G*01:03	0,00	1,00
A*25:01 G*01:03	0,00	1,00
A*01:01 G*01:06	0,03	1,00
<b>A*30:01 G*01:05N</b>	<b>0,01</b>	<b>0,66</b>
<b>A*31:01 G*01:05N</b>	<b>0,00</b>	<b>0,31</b>

n; número de individuos.

D; valor de desequilibrio de ligamiento

D'; valor de desequilibrio de ligamiento estandarizado.

En las tablas R4 y R5, se aprecia como los dos haplotipos biloci más frecuentes en esta población, HLA-A\*02:01-G\*01:01 (24,6%) y HLA-A\*24:02-G\*01:04 (13,2%), también se encuentran, sus alelos, en desequilibrio de ligamiento,  $D=0,07$ ;  $D'=0,94$  y  $D=0,10$ ;  $D'=1,00$  respectivamente para cada haplotipo.

La distancia física entre los loci HLA-A y HLA-G, parece no ser superior a las 300 kb (Morales et al., 1993), así el sorprendente nivel de desequilibrio de ligamiento entre los alelos de ambos loci, es muy similar al observado en la región de clase II entre los loci, HLA-DRB1 y HLA-DQB1, donde evidencias frecuentes de recombinación no han sido observadas (Ober et al., 1996). Corroborando lo observado en estudios anteriores (Morales et al., 1993), se muestra en la tabla R5 como el alelo HLA-G\*01:01 está en desequilibrio de ligamiento con el grupo de alelos HLA-A\*02, -A\*03 y -A\*29, y el alelo HLA-G\*01:04 con el grupo de alelos HLA-A\*23 y -A\*24. Es de resaltar como estos desequilibrios de ligamiento son, por tanto, similares en caucásicos y amerindios.

Respecto al alelo HLA-G\*01:05N, en la tabla R4 se aprecia que la frecuencia de los haplotipos biloci que llevan este alelo, HLA-A\*30:01-G\*01:05N y HLA-A\*31:01-G\*01:05N, es bastante baja, 0,8% y 0,4%, respectivamente, ambas por debajo del 1%. Sin embargo, en la tabla R5 aparece un valor positivo para el desequilibrio de ligamiento entre HLA-G\*01:05N-HLA-A\*30:01 ( $D=0,01$ ;  $D'=0,66$ )

### 3.2 Loci HLA-B y -G.

Las frecuencias del haplotipo biloci HLA-B-G se han calculado gracias al programa Arlequín v2.0 (Schneider et al., 2000). En la tabla R6 se muestran los diez haplotipos más frecuentes en la población de estudio, más las frecuencias de los haplotipos que incluyen el alelo nulo HLA-G\*01:05N

**Tabla R6:** Frecuencias haplotípicas de los loci HLA-B y HLA-G, con cuatro dígitos de resolución, en amerindios inmigrantes en Madrid. (n=89). En negrita se marca marcan los haplotipos que porten el alelo nulo.

Loci HLA-B y -G	Frecuencias %
B*44:03 G*01:01	6,2
B*51:01 G*01:01	5,5
B*40:02 G*01:01	4,5
B*48:01 G*01:01	3,9
B*35:01 G*01:01	3,9
B*40:04 G*01:01	3,4
B*35:04 G*01:01	3,4
B*07:02 G*01:01	3,2
B*35:01 G*01:04	2,8
B*38:01 G*01:01	2,2
<b>B*15:05 G*01:05N</b>	<b>1,1</b>
<b>B*57:03 G*01:05N</b>	<b>0,5</b>

n; número de individuos.

El cálculo del valor de desequilibrio de ligamiento entre estos dos loci, se ha realizado gracias al programa Arlequín v3.0 (Excoffier et al., 2005). Diez haplotipos se han encontrado cuyos alelos están es desequilibrio de ligamiento positivo ( $D>0$ ) y  $p<0.05$ .

**Tabla R7:** Desequilibrio de ligamiento entre los alelos de los haplotipos biloci HLA-B y HLA-G, con cuatro dígitos de resolución, en amerindios inmigrantes en Madrid. (n=89). En negrita se marcan los valores de desequilibrio de ligamiento positivos que porten el alelo nulo.

Loci HLA-B y -G	D	D'
B*44:03 G*01:01	0,02	1,00
B*35:01 G*01:04	0,02	0,37
B*15:03 G*01:04	0,01	1,00
B*58:02 G*01:04	0,00	1,00
B*18:01 G*01:04	0,01	1,00
B*35:43 G*01:04	0,00	1,00
B*39:03 G*01:04	0,00	1,00
B*50:01 G*01:04	0,00	1,00
B*39:06 G*01:03	0,01	0,30
B*14:02 G*01:03	0,01	0,48
B*39:01 G*01:03	0,00	0,30
B*27:05 G*01:03	0,00	0,30
B*08:12 G*01:06	0,00	1,00
B*15:16 G*01:06	0,00	1,00
B*57:01 G*01:06	0,01	0,37
B*37:01 G*01:06	0,00	1,00
<b>B*15:05 G*01:05N</b>	<b>0,01</b>	<b>1,00</b>
<b>B*57:03 G*01:05N</b>	<b>0,01</b>	<b>1,00</b>

*n*; número de individuos.

*D*; valor de desequilibrio de ligamiento

*D'*; valor de desequilibrio de ligamiento estandarizado.

Se observa como el alelo más frecuente de HLA-G, G\*01:01 (71,8% con cuatro dígitos de resolución) se encuentra en desequilibrio de ligamiento con el alelo HLA-B\*44:03 ( $D=0,02$ ;  $D'=1,00$ ), y el segundo alelo más frecuente, HLA-G\*01:04 (18,5% con cuatro dígitos de resolución), sin embargo, se encuentra en desequilibrio de ligamiento con siete alelos HLA-B diferentes (ver tabla R7). Por otro lado, el alelo nulo, G\*01:05N, se encuentra en desequilibrio de ligamiento con el alelo HLA-B\*15:05 ( $D=0,01$ ;  $D'=1,00$ ) y con HLA-B\*57:03 ( $D=0,01$ ;  $D'=1,00$ ).

### 3.3 Haplotipo extendido HLA-A-B-G

Las frecuencias del haplotipo extendido HLA-A-B-G, se han calculado gracias al programa Arlequín v2.0 (Schneider et al., 2000). En la tabla R8 se muestran los ocho haplotipos más frecuentes en la población de estudio, más las frecuencias de los haplotipos que incluyen el alelo nulo HLA-G\*01:05N.

**Tabla R8:** Frecuencias haplotípicas del haplotipo extendido HLA-A-DRB1-G, con cuatro dígitos de resolución, en amerindios inmigrantes en Madrid. (n=83). En negrita se marca marcan los haplotipos que porten el alelo nulo.

Haplotipos extendidos HLA-A-B-G	Frecuencias %
A*02:01 DRB1*08:02 G*01:01	3
A*02:01 DRB1*04:07 G*01:01	3
A*02:01 DRB1*09:01 G*01:01	2,4
A*24:02 DRB1*14:02 G*01:04	2,4
A*02:01 DRB1*04:03 G*01:01	2,4
A*24:02 DRB1*07:01 G*01:04	2,4
A*02:01 DRB1*03:01 G*01:01	2,4
A*03:01 DRB1*15:01 G*01:01	2,4
A*26:01 DRB1*01:01 G*01:01	1,7
A*68:01 DRB1*04:07 G*01:01	1,7
A*02:01 DRB1*14:02 G*01:01	1,3
<b>A*30:01 DRB1*04:03 G*01:05N</b>	<b>0,4</b>
<b>A*31:01 DRB1*01:01 G*01:05N</b>	<b>0,4</b>
<b>A*33:03 DRB1*13:02 G*01:05N</b>	<b>0,4</b>

n; número de individuos.

En la tabla R8 se muestra como los haplotipos extendidos HLA-A-B-G más frecuentes en la población de estudio son el HLA-A\*24:02-B\*35:01-G\*01:04 y el HLA-A\*29:02-B\*44:03-G\*01:01, ambos con una frecuencia del 3,0%. El alelo HLA-A\*24:02 es uno de los alelos fundadores de clase I en la población amerindia.

### 3.4 Haplotipo extendido HLA-A-DRB1-G.

Las frecuencias del haplotipo extendido HLA-A-DRB1-G, se han calculado gracias al programa Arlequín v2.0 (Schneider et al., 2000). Se muestran los once haplotipos más frecuentes en la población de estudio, más las frecuencias de los haplotipos que incluyen el alelo nulo HLA-G\*01:05N.

**Tabla R9:** Frecuencias haplotípicas del haplotipo extendido HLA-A-DRB1-G, con cuatro dígitos de resolución, en amerindios inmigrantes en Madrid. (n=116). En negrita se marca marcan los haplotipos que porten el alelo nulo.



Haplotipos extendidos HLA-A-DRB1-G	Frecuencias %
A*02:01-DRB1*08:02-G*01:01	4,7
A*02:01-DRB1*04:07-G*01:01	4,2
A*02:01-DRB1*09:01-G*01:01	3,0
A*24:02-DRB1*14:02-G*01:04	2,9
A*02:01-DRB1*04:03-G*01:01	2,2
A*24:02-DRB1*07:01-G*01:04	2,2
A*02:01-DRB1*03:01-G*01:01	1,7
A*03:01-DRB1*15:01-G*01:01	1,7
A*26:01-DRB1*01:01-G*01:01	1,7
A*68:01-DRB1*04:07-G*01:01	1,7
A*02:01-DRB1*14:02-G*01:01	1,3
<b>A*30:01-DRB1*04:03-G*01:05N</b>	<b>0,4</b>
<b>A*31:01-DRB1*01:01-G*01:05N</b>	<b>0,4</b>
<b>A*33:03-DRB1*13:02-G*01:05N</b>	<b>0,4</b>

*n*; número de individuos.

Se puede observar como los dos haplotipos más frecuentes en la población ameríndia inmigrante en Madrid son, HLA-A\*02:01-DRB1\*08:02-G\*01:01 y HLA-A\*02:01-DRB1\*04:07-G\*01:01 con unas frecuencias de 4,7 y 4,2 % respectivamente.

En este caso, el alelo HLA-A\*02:01 también forma parte del los alelos fundadores en la población ameríndia. Respecto al locus DRB1, el alelo -DRB1\*04:07 también es propio de esta población.

### 3.5 Haplotipo extendido HLA-A-B-DRB1-G.

De nuevo, las frecuencias del haplotipo extendido HLA-A-B-DRB1-G, se han calculado gracias al programa Arlequín v2.0 (Schneider et al., 2000). Se muestran los nueve haplotipos más frecuentes en la población de estudio, más las frecuencias de los haplotipos que incluyen el alelo nulo HLA-G\*01:05N.

**Tabla R10:** Frecuencias haplotípicas del haplotipo extendido HLA-A -B -DRB1 -G, con cuatro dígitos de resolución, en amerindios inmigrantes en Madrid. (n=86). En negrita se marcan los haplotipos que porten el alelo nulo.

Haplotipos HLA-A-B-DRB1-G	Frecuencias %	Posible origen
A*24:02-B*35:01-DRB1*14:02-G*01:04 <sup>a</sup>	2,9	amerindio
A*02:01-B*35:04-DRB1*08:02-G*01:01 <sup>b</sup>	1,7	amerindio
A*02:01-B*40:02-DRB1*04:07-G*01:01 <sup>c</sup>	1,2	amerindio
A*03:01-B*07:02-DRB1*15:01-G*01:01 <sup>d</sup>	1,2	europeo
A*02:01-B*44:03-DRB1*13:02-G*01:01 <sup>e</sup>	1,2	desconocido
A*03:01-B*51:01-DRB1*04:04-G*01:01 <sup>f</sup>	1,2	desconocido
A*24:02-B*14:02-DRB1*07:01-G*01:01 <sup>g</sup>	1,2	desconocido
A*68:01-B*40:04-DRB1*04:03-G*01:01 <sup>h</sup>	1,2	desconocido
A*31:01-B*39:06-DRB1*13:02-G*01:01 <sup>i</sup>	1,2	nuevo
<b>A*29:02-B*15:05-DRB1*04:03-G*01:05N<sup>j</sup></b>	<b>0,6</b>	nuevo
<b>A*31:01-B*15:05-DRB1*04:03-G*01:05N<sup>k</sup></b>	<b>0,6</b>	nuevo
<b>A*33:03-B*57:03-DRB1*13:02-G*01:05N<sup>l</sup></b>	<b>0,6</b>	nuevo

n; número de individuos.

a; este haplotipo, sin HLA-G, también se encuentra en aymaras, quechuas y jaidukama, (Arnaiz-Villena et al., 2005; Martínez-Laso et al., 2006; Martínez-Laso et al., 2011).

b; este haplotipo, sin HLA-G, también se encuentra en aymaras, mayos, mayas, quechuas, wayus y nahuas (Arnaiz-Villena et al., 2005; Arnaiz-Villena et al., 2007; Gomez-Casado et al., 2003; Martínez-Laso et al., 2006; Silvera et al., 2011; Vargas-Alarcón et al., 2007).

c; este haplotipo, sin HLA-G, también se encuentra en aymaras, mayos, jaidukama y lamas. (Arnaiz-Villena et al., 2005; Arnaiz-Villena et al., 2007; Martínez-Laso et al., 2011; Moscoso et al., 2006).

d; este haplotipo, sin HLA-G, también se encuentra en poblaciones de Europa occidental, chuvassios etc. (Arnaiz-Villena et al., 2003; Sánchez-Velasco et al., 2003).

e; este haplotipo se ha encontrado en frecuencias inferiores a 0,01%. (Gonzalez-Galarza et al., 2011)

f; este haplotipo se ha encontrado en frecuencias inferiores a 0,01%. (Gonzalez-Galarza et al., 2011)

g; este haplotipo se ha encontrado en frecuencias inferiores a 0,01%. (Gonzalez-Galarza et al., 2011)

h; este haplotipo se ha encontrado en frecuencias inferiores a 0,01%. (Gonzalez-Galarza et al., 2011)

i; este haplotipo no se ha encontrado en ninguna otra población.

j; este haplotipo no se ha encontrado en ninguna otra población.

k; este haplotipo no se ha encontrado en ninguna otra población.

l; este haplotipo no se ha encontrado en ninguna otra población.

El haplotipo extendido más frecuente en la población de estudio es el HLA-A\*24:02-B\*35:01-DRB1\*14:02-G\*01:04, con una frecuencia que roza el 3% de la población. Este haplotipo, sin el locus HLA-G, también se encuentra en otras poblaciones amerindias como los aymaras (2,8%) (Arnaiz-Villena et al., 2005), quechuas (1,4%) (Martínez-Laso et al., 2006) y jaidukama (Martínez-Laso et al., 2011). Los alelos de este haplotipo, HLA-A\*24:02 y HLA-G\*01:04, se encuentran en desequilibrio de ligamiento ( $D=0,10$ ;  $D'=1,00$ ) en esta población.

El segundo haplotipo extendido más frecuente con casi un 2%, es el HLA-A\*02:01-B\*35:04-DRB1\*08:02-G\*01:01. Este, sin el locus HLA-G, también se puede encontrar

en otras poblaciones amerindias como, mayas (8,4%)(Gomez-Casado et al., 2003), aymaras (10,4%)(Arnaiz-Villena et al., 2005), quechuas (7,5%)(Martínez-Laso et al., 2006), mayos (2,5%)(Arnaiz-Villena et al., 2007), nahuas (6,1%)(Vargas-Alarcón et al., 2007) y wayus (3,2%)(Silvera et al., 2011). Los alelos de este haplotipo, HLA-A\*02:01 y HLA-G\*01:01, se encuentran en desequilibrio de ligamiento ( $D=0,07$ ;  $D'=0,94$ ) en esta población (ver tabla R5).

Es conveniente reseñar, que en los párrafos anteriores no se ha tenido en cuenta el locus HLA-G, en relación a la comparativa de los haplotipos extendidos con otras poblaciones mundiales, por la falta de estudios que lo incluyan en los mismos.

Se puede observar como los haplotipos extendidos más frecuentes (sin incluir el locus HLA-G), tienen un posible origen amerindio. Existe una cierta influencia europea, que se puede observar en el haplotipo HLA-A\*03:01-B\*07:02-DRB1\*15:01-G\*01:01, dado que tiene un posible origen norte africano/ europeo occidental y tiene una alta frecuencia en poblaciones del norte de España (Sánchez-Velasco et al., 2003), Irlanda, por ejemplo (Gonzalez-Galarza et al., 2011).

Todos los haplotipos extendidos llevan el alelo HLA-G\*01:01 exceptuando el más frecuente de ellos, que lleva el alelo HLA-G\*01:04 (segundo más frecuente en la población de estudio, 18,5%).

Los haplotipos del *e* al *h*, no se han encontrado en ninguna otra población con frecuencia mayor al 0,5% y el haplotipo *i*, no se ha encontrado en ninguna otra población independientemente a la frecuencia, por lo que podría ser un haplotipo extendido nuevo.

Los haplotipos extendidos que llevan el alelo HLA-G\*01:05N se comentan en el apartado 4.

#### **4. ESTUDIO DEL ALELO NULO, HLA-G\*01:05N, EN LA POBLACIÓN AMERINDIA INMIGRANTE EN MADRID.**

##### **4.1 Descripción alelos nulos encontrados en la muestra.**

Como se puede observar en el listado de alelos HLA-G determinados en la población de estudio (ver anexoVI), existen tres alelos nulos encontrados en tres individuos diferentes en heterocigosis. Cada uno de estos tres individuos, procede de un

país de origen distinto, Colombia, Argentina y Nicaragua, dos hombres y una mujer respectivamente.

#### **4.2 Frecuencias de los alelos nulos encontrados en la muestra.**

Como ya se ha descrito en el apartado 2 anterior, las frecuencias de estos alelos en la población amerindia inmigrante en Madrid, son de:

- HLA-G\*01:05N: 1,0% con cuatro dígitos de resolución, (n=154), (ver tabla R2)
- HLA-G\*01:05N: 1,3% con seis dígitos de resolución, (n=117), (ver tabla R3)

La frecuencia del alelo nulo HLA-G\*01:05N, es de solo 1,0-1,3%, con cuatro y seis dígitos de resolución respectivamente, es decir, una frecuencia baja. Esto mismo ocurre en poblaciones caucásicas europeas, orientales, en pigmeos africanos y brasileños. En el caso de las poblaciones amerindias, mayas y uros, la presencia de este alelo es nula (Arnaiz-Villena et al., 2013). Sin embargo, las mayores frecuencias del alelo HLA-G\*10:05N, se han encontrado en poblaciones de Oriente Medio y en negros africanos. (Arnaiz-Villena et al., 2013).

#### **4.3 El alelo nulo y los haplotipos extendidos.**

Como se observa en la tabla R10 del apartado 3 anterior, las frecuencias de los haplotipos extendidos que llevan este alelo, en la población amerindia inmigrante en Madrid, son:

- HLA-A\*33:03-B\*57:03-DRB1\*13:02-G\*01:05N: 0,6%, (n=86)
- HLA-A\*31:01-B\*15:05-DRB1\*04:03-G\*01:05N: 0,6%, (n=86)
- HLA-A\*29:02-B\*15:05-DRB1\*04:03-G\*01:05N: 0,6%, (n=86)

Estos tres haplotipos, presentan la misma frecuencia (0,6%), que es la más baja encontrada en la población de estudio. Por otro lado, ninguno de ellos se ha encontrado en ninguna otra población.

#### 4.4 Comparación de las frecuencias del alelo nulo de la población de estudio, con otras poblaciones mundiales estudiadas.

Para comparar las frecuencias entre las distintas poblaciones que han estudiado este alelo y la población amerindia inmigrante en Madrid, se ha usado el Test de Fisher del programa estadístico SPSS v.17.0 (versión 17.0 para Windows; SPSS, Inc., Chicago, IL, USA). La significación estadística ha sido definida como  $p < 0.05$ .

Se ha comparado la frecuencia de HLA-G\*01:05N, entre la población de estudio y diecisiete poblaciones de origen caucásico, amerindio, oriental o africano.

**Tabla R11:** Diferencias en la frecuencia del alelo HLA-G\*01:05N entre amerindios inmigrantes en Madrid y otras poblaciones mundiales.

Etnicidad	Población	n	HLA-G*01:05N %	p-valor <sup>a</sup>	Referencias
<i>Amerindios</i>	Amerindios inmigrantes en Madrid	154	1,0	-	-
	Brasileños	103	1,0	n.s.	(Castelli et al., 2007)
	Uros	105	0,0	n.s.	(Arnaiz-Villena et al., 2013)
	Mayas	132	0,0	n.s.	(Arnaiz-Villena et al., 2013)
<i>Caucásicos</i>	Indios del Norte de India	120	15,4	0	(Abbas et al., 2004)
	Iraníes	102	18,1	0	(Rahimi et al., 2010)
	Iraquíes	97	8,2	0	(Jassem et al., 2012)
	Europeos	52	1,0	n.s.	(Hviid, 2006)
<i>Europeos</i>	Espanoles	114	3,0	n.s.	(Suarez et al., 1997)
	Alemanes	132	2,3	n.s.	(van der Ven et al., 1998)
	Franceses	57	1,8	n.s.	(Di Cristofaro et al., 2011)
	Polacos	100	1,5	n.s.	(Sipak-Szmigiel et al., 2009)
<i>Orientales</i>	Chinos Han	146	1,4	n.s.	(Yan et al., 2006)
	Coreanos	200	2,3	n.s.	(Park et al., 2012)
<i>Africanos</i>	Africanos de Ghana	42	4,8	0,041	(Ishitani et al., 1999)
	Afroamericanos	42	8,3	0,001	(Ishitani et al., 1999)
	Pigmeos	74	1,4	n.s.	(Di Cristofaro et al., 2011)
	Africanos de Shona	108	11,1	0	(Matte et al., 2000)

n; número de individuos de la muestra.

a; Test de Fisher (dos-colas) entre las frecuencias de HLA-G\*01:05N de amerindios inmigrantes en Madrid y otras poblaciones mundiales.

n.s.; no significativo.

Según los datos mostrados en la tabla R11 y figura R3, las poblaciones amerindias (brasileños, uros, mayas y amerindios inmigrantes en Madrid) presentan las frecuencias más bajas de este alelo. Mientras que las frecuencias más elevadas, se observan en las poblaciones del Oriente Medio (indios del Norte de India, iraníes e iraquíes) y las poblaciones africanas, excepto pigmeos, (africanos de Ghana, afroamericano y africanos de Shona).

Estas seis poblaciones, cuyas frecuencias del alelo nulo son más elevadas, son las únicas que presentan diferencias significativas con  $p < 0.05$ , con los amerindios inmigrantes en Madrid.

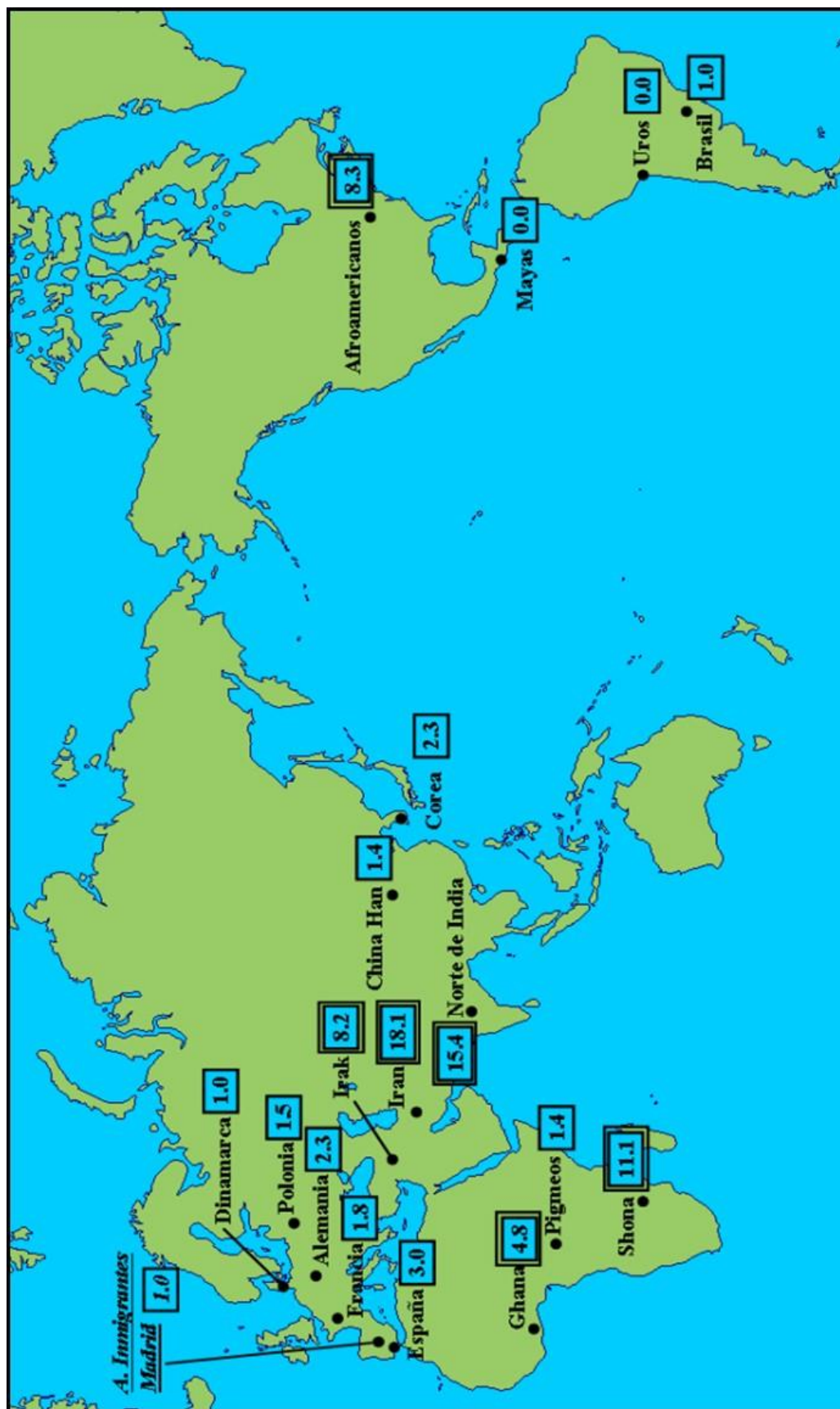


Figura R3: Mapa de las poblaciones mundiales utilizadas para comparar las frecuencias alélicas de HLA-G\*01:05N, con la población de estudio, amerindios inmigrantes en Madrid. En los recuadros se muestran las frecuencias de cada población, con doble recuadro se marcan las poblaciones con mayor frecuencia para este alelo.

#### 4.5 Alelo nulo y el desequilibrio de ligamiento con el locus HLA-A.

Como se ha mencionado con anterioridad, el locus HLA-G se encuentra en fuerte desequilibrio de ligamiento con el locus HLA-A, debido a su cercanía física dentro del cromosoma (Ober et al., 1996). Por ello, se han calculado los valores de desequilibrio de ligamiento entre los alelos de ambos loci en la población de estudio, incluyendo los alelos nulos presentes.

- HLA-A\*30:01 G\*01:05N:  $D=0,01$ ;  $D'=0,66$
- HLA-A\*31:01 G\*01:05N:  $D=0,00$ ;  $D'=0,31$

Corroborando lo expuesto en estudios anteriores (Arnaiz-Villena et al., 2013) el alelo nulo HLA-G\*01:05N se encuentra en desequilibrio de ligamiento con el alelo HLA-A\*30:01, también en la población amerindia inmigrante en Madrid.

### **5. ESTUDIO DE LAS NUEVAS PROTEÍNAS HLA-G ENCONTRADAS EN LA POBLACIÓN AMERINDIA INMIGRANTE EN MADRID.**

En el proceso de determinación de los alelos HLA-G en la población de estudio, se han encontrado dos alelos nuevos, en dos individuos de diferente origen. El primero de ellos, de origen ecuatoriano (alelo nuevo 1) y el segundo (alelo nuevo 2), de origen dominicano.

#### **5.1 Secuencias y estructuras de los nuevos alelos.**

##### **5.1.1 Un alelo nuevo en los Andes, Ecuador. HLA-G\*01:08:02.**

##### **• Secuencia de ADN:**

El nombre oficial asignado por el Comité de Nomenclatura (Robinson et al., 2015) ([http://hla.alleles.org/data/refs/g\\_refs.html](http://hla.alleles.org/data/refs/g_refs.html)) es HLA-G\*01:08:02 (Arnaiz-Villena et al., 2016).

A continuación, se muestra la secuencia del alelo nuevo.



ADN	74	
		<b>Exón 2</b>
G*01:01:01:01		GC TCC CAC TCC ATG AGG TAT TTC AGC GCC GCC GTG TCC CGG CCC GGC CGC GGG GAG CCC CGC TTC ATC GCC ATG GGC
Alelo nuevo 1		-----
	151	
G*01:01:01:01		TAC GTG GAC GAC ACG CAG TTC GTG CGG TTC GAC AGC GAC TCG GCG TGT CCG AGG ATG GAG CCG CGG CCG TGG
Alelo nuevo 1		-----
	226	
G*01:01:01:01		GTG GAG CAG GAG GGG CCG GAG TAT TGG GAA GAG GAG ACA CGG AAC ACC AAG GCC CAC GCA CAG ACT GAC AGA ATG
Alelo nuevo 1		-----A-----
	301	
		<b>Exón 3</b>
G*01:01:01:01		AAC CTG CAG ACC CTG CGC GGC TAC TAC AAC CAG AGC GAG GCC A GT TCT CAC ACC CTC CAG TGG ATG ATT GGC TGC
Alelo nuevo 1		-----T-----
	376	
G*01:01:01:01		GAC CTG GGG TCC GAC GGA CGC CTC CTC CGC GGG TAT GAA CAG TAT GCC TAC GAT GGC AAG GAT TAC CTC GCC CTG
Alelo nuevo 1		-----
	451	
G*01:01:01:01		AAC GAG GAC CTG CGC TCC TGG ACC GCA GCG GAC ACT GCG GCT CAG ATC TCC AAG CGC AAG TGT GAG GCG GCC AAT
Alelo nuevo 1		-----
	526	
G*01:01:01:01		GTG GCT GAA CAA AGG AGA GCC TAC CTG GAG GGC ACG TGC GTG GAG TGG CTC CAC AGA TAC CTG GAG AAC GGG AAG
Alelo nuevo 1		-----

Continuación:	
	<b>Exón 4</b>
601	
G*01:01:01:01	GAG ATG CTG CAG CGC GCG G AC CCC CCC AAG ACA CAC GTG ACC CAC CAC CCT GTC TTT GAC TAT GAG GCC ACC CTG
Alelo nuevo 1	----- -----
676	
G*01:01:01:01	AGG TGC TGG GCC CTG GGC TTC TAC CCT GCG GAG ATC ATA CTG ACC TGG CAG CGG GAT GGG GAG GAC CAG ACC CAG
Alelo nuevo 1	-----T-----
751	
G*01:01:01:01	GAC GTG GAG CTC GTG GAG ACC AGG CCT GCA GGG GAT GGA ACC TTC CAG AAG TGG GCA GCT GTG GTG CCT TCT
Alelo nuevo 1	-----
826	
G*01:01:01:01	GGA GAG GAG CAG AGA TAC ACG TGC CAT GTG CAG CAT GAG GGG CTG CCG GAG CCC CTC ATG CTG AGA TGG A
Alelo nuevo 1	----- -----

Figura R4: Secuencia del alelo nuevo 1 (individuo de Ecuador), comparada con una secuencia de referencia (alelo HLA-G\*01:01:01), sin intrones. Se indica con guiones (-) los nucleótidos iguales a la secuencia de referencia. Las barras verticales (|) indican fin de un exón. El número al inicio de cada línea, indica el número de posición de ese nucleótido en la secuencia. Ver correspondencia de letras en Anexos II.

En la figura R4 se muestra la secuencia de ADN nueva del individuo de origen ecuatoriano, de sexo masculino y procedente de un área rural, tanto él como sus abuelos y antepasados registrados.

Para un mejor análisis de la misma, se ha comparado con una secuencia de referencia (HLA-G\*01:01:01:01), sin introducir intrones en ninguna.

Para esta secuencia nueva, se han determinado los exones 2, 3 y 4. A lo largo de estos tres exones, se pueden localizar cuatro posiciones donde el nucleótido cambia, creando una secuencia nueva no vista en ningún individuo anteriormente. En cada una de estas cuatro posiciones hay un SNP ya descrito, pero no la combinación de todos ellos. Estos, se reparten a lo largo de los tres exones, un SNP-R en el exón 2 (el codón pasa de CCG a CCA), dos SNPs-Y (el primer codón pasa de CAC a CAT y el segundo de GGC a GGT), ambos en el exón 3, y un SNP-Y en el exón 4 (el codón pasa de CGG a TGG) (ver correspondencia en anexo II). En concreto, el alelo nuevo consta de: una adenina en el SNP-R, y tres timinas en los tres SNPs-Y restantes (ver figura R4). Esto no indica que no puedan existir más cambios en el resto de exones.

- **Secuencia aminoacídica:**

Se ha determinado la secuencia aminoacídica de dicho alelo nuevo, gracias al programa MEGA 7 (Kumar et al., 2015).

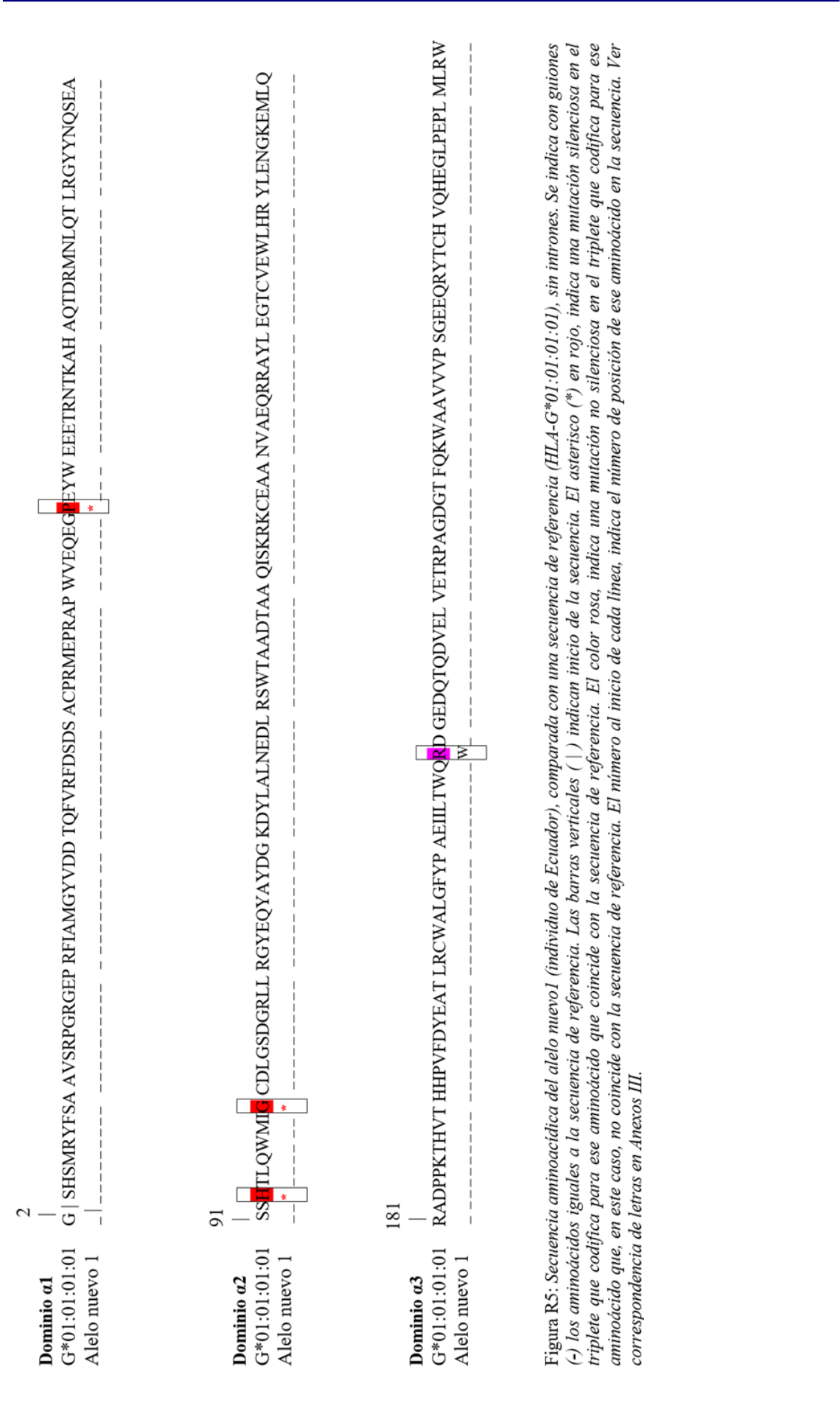


Figura R5: Secuencia aminoacídica del alelo nuevo1 (individuo de Ecuador), comparada con una secuencia de referencia (HLA-G\*01:01:01), sin intrones. Se indica con guiones (-) los aminoácidos iguales a la secuencia de referencia. Las barras verticales (|) indican inicio de la secuencia. El asterisco (\*) en rojo, indica una mutación silenciosa en el triplete que codifica para ese aminoácido que coincide con la secuencia de referencia. El color rosa, indica una mutación no silenciosa en el triplete que codifica para ese aminoácido que, en este caso, no coincide con la secuencia de referencia. El número al inicio de cada línea, indica el número de ese aminoácido en la secuencia. Ver correspondencia de letras en Anexos III.

	Dominio $\alpha 1$	Dominio $\alpha 2$	Dominio $\alpha 3$
Residuo	57	93	100
G*01:01:01	Pro	His	Gly
Alelo nuevo Ecuador	Pro	His	Gly
Cambios en codones ADN	CCG→CCA	CAC→CAT	GGC→GGT
			CGG→TGG

Figura R6: Variabilidad de aminoácidos entre el alelo nuevo de Ecuador y el alelo HLA-G\*01:01. La secuencia de este alelo fue enviada al GenBank el 29 de enero del 2015; el número de acceso KP739973 fue dado el 4 de febrero del 2015. El consenso de aa está basado en la secuencia aminoacídica de HLA-G\*01:01. La mutación silenciosa que hace a este alelo nuevo, está marcada con un recuadro, de otro modo sería igual al alelo HLA-G\*01:08.

	Exon 2					Exon 3					Exon 4				
	13	27	31	34	54	57	93	100	104	105	110	159	169	171	219
HLA-G*01:08:02 <sup>a</sup>						CCA	CAT	GGT							TGG
															Trp
HLA-G*01:01	TCC	TAC	ACG	GTG	CAG	CCG	CAC	GGC	GGG	TCC	CTC	TAC	CAC	TAC	CGG
	Ser	Tyr	Thr	Val	Gln	Pro	His	Gly	Gly	Ser	Leu	Tyr	His	Tyr	Arg
HLA-G*01:02					CGG										
					Arg										
HLA-G*01:03			TCG												
			Ser												
HLA-G*01:04											ATC				
											Ile				
HLA-G*01:06															
HLA-G*01:07	TTC										ATC				
	Phe										Ile				
HLA-G*01:08															TGG
															Trp
HLA-G*01:09												CAC			
												His			
HLA-G*01:10			ATG												
			Met												
HLA-G*01:11			ATG								ATC				
			Met								Ile				
HLA-G*01:12		CAC													
		His													
HLA-G*01:14								GAC							
								Asp							
HLA-G*01:15									GTG		ATC				
									Val		Ile				
HLA-G*01:16										TGC					
										Cys					
HLA-G*01:17												CGC	CAC		
												Arg	His		
HLA-G*01:18															
HLA-G*01:19			CTG								ATC				
			Val								Ile				

<sup>a</sup>HLA-G\*01:08:02 (alelo nuevo 1 de Ecuador): es una variante genética de HLA-G\*01:08. Solo se indican los codones, mutaciones sinónimas y los aa codificados.

Figura R7: Proteínas de los alelos HLA-G. Se incluyen las posiciones variables de los aa de cada alelo HLA-G.

En la figura R5 se muestra la secuencia de aa nueva del individuo de origen ecuatoriano. Para un mejor análisis de la misma, se ha comparado con una secuencia de referencia (HLA-G\*01:01:01:01), sin introducir intrones en ninguna.

De tal manera, para el alelo nuevo 1 de Ecuador, al tener las secuencias de sus exones 2, 3 y 4, se ha podido obtener las secuencias de aa de los dominios  $\alpha 1$  (codificada por exón 2),  $\alpha 2$  (codificada por exón 3) y  $\alpha 3$  (codificada por exón 4). En la figura R4, se muestra como los cambios se producen en los tres exones, por ello, en la proteína los cambios se podrían observar a lo largo de los tres dominios. Sin embargo, puede ocurrir lo que se conoce como mutaciones o sustituciones silenciosas, aquellas en las que el cambio de nucleótido no produce, en ese triplete, un cambio de aa, por lo que, el resultado en la proteína es el mismo que si no se hubiese producido dicha sustitución. En este caso, de los cuatro cambios de nucleótidos que hay en la secuencia, el residuo 57: prolina→prolina, residuo 93: histidina→histidina y residuo 100: glicina→glicina, no producen cambio de aa en la proteína, como se observa en la figura R6. No es el caso del último cambio observado en el exón 4, donde hay un SNP-Y, es decir, puede ser una timina o una citosina (ver figura R4). En el caso de este alelo, es una timina que se encuentra situada en la primera posición de un triplete que codifica para triptófano (W) (TGG), por lo que el residuo 219 pasa de arginina a triptófano (ver figura R4). Esta posición se conoce como un sitio común de polimorfismo en el dominio  $\alpha 3$  de HLA-G, y está relacionada con una posible diversidad de afinidad entre este dominio y CD8 $\alpha\alpha$  y LILRBs (Donadi et al., 2011; Gao et al., 2000). Para observar con mayor claridad los cambios producidos en la nueva proteína (alelo nuevo 1 de Ecuador), en la figura R7 se esquematizan dichos cambios comparándolos con el resto de proteínas de HLA-G.

El haplotipo extendido de este individuo es: HLA-A\* 29:01,\* 32:01; HLA-B\* 14:01,\* 44:03; HLA-DRB1\*07:01,\* 07:01; HLA-DQB1\*02:01,\* 02:01; HLA-G\*01:01:01.

### 5.1.2 Un alelo nuevo en el Caribe, República Dominicana. HLA-G\*01:20

#### • Secuencia de ADN:

El nombre oficial asignado por el Comité de Nomenclatura (Robinson et al., 2015) ([http://hla.alleles.org/data/refs/g\\_refs.html](http://hla.alleles.org/data/refs/g_refs.html)) es HLA-G\*01:20 (Arnaiz-Villena et al., 2016).

A continuación, se muestra la secuencia del alelo nuevo.

ADN	74	
		<b>Exón 2</b>
G*01:01:01:01		GC TCC CAC TCC ATG AGG TAT TTC AGC GCC GCC GTG TCC CGG CCC GGC CGC GGG GAG CCC CGC TTC ATC GCC ATG GGC
Alelo nuevo 2		-----
	151	
G*01:01:01:01:01		TAC GTG GAC GAC ACG CAG TTC GTG CGG TTC GAC AGC GAC TCG GCG TGT CCG AGG ATG GAG CCG CGG GCG CCG TGG
Alelo nuevo 2		-----
	226	
G*01:01:01:01:01		GTG GAG CAG GAG GGG CCG GAG TAT TGG GAA GAG GAG ACA CGG AAC ACC AAG GCC CAC GCA CAG ACT GAC AGA ATG
Alelo nuevo 2		-----
	301	
		<b>Exón 3</b>
G*01:01:01:01:01		AAC CTG CAG ACC CTG CGC GGC TAC TAC AAC CAG AGC GAG GCC A GT TCT CAC ACC CTC CAG TGG ATG ATT GGC TGC
Alelo nuevo 2		-----
	376	
G*01:01:01:01:01		GAC CTG GGG TCC GAC GGA CGC CTC CTC CGC GGG TAT GAA CAG TAT GCC TAC GAT GGC AAG GAT TAC CTC GCC CTG
Alelo nuevo 2		-----T-----
	451	
G*01:01:01:01:01		AAC GAG GAC CTG CGC TCC TGG ACC GCA GCG GAC ACT GCG GCT CAG ATC TCC AAG CGC AAG TGT GAG GCG GCC AAT
Alelo nuevo 2		-----



Continuación:	
	526
G*01:01:01:01	GTG GCT GAA CAA AGG AGA GCC TAC CTG GAG GGC ACG TGC GTG GAG TGG CTC CAC AGA TAC CTG GAG AAC GGG AAG
Alelo nuevo 2	-----
	601
G*01:01:01:01	GAG ATG CTG CAG CGC GCG G
Alelo nuevo 2	---- -C_ -----

Figura R8: Secuencia del alelo nuevo 2 (individuo de R. Dominicana), comparada con una secuencia de referencia (alelo HLA-G\*01:01:01), sin intrones. Se indica con guiones (-) los nucleótidos iguales a la secuencia de referencia. Las barras verticales (|) indican fin de un exón. El número al inicio de cada línea, indica el número de posición de ese nucleótido en la secuencia. Ver correspondencia de letras en Anexos II.



El alelo nuevo 2 de República Dominicana, pertenece a un individuo de sexo femenino y que únicamente permaneció en Madrid dos semanas, ella también procede de un área rural, así como sus abuelos y antecesores registrados.

A pesar de que es difícil encontrar características amerindias en islas del Caribe, desde el punto de vista fenotípico, se ha demostrado que alrededor de un 15% de los alelos HLA proceden un origen amerindio (Alegre et al., 2007).

Se han determinado los exones 2 y 3. Las variaciones nucleotídicas localizadas en este alelo, se encuentran únicamente en el exón 3. De tal manera, se observan dos SNPs, en primer lugar, un SNP-W y en segundo lugar un SNP-Y (ver correspondencia en anexo II). En este caso, el primer SNP ya estaba descrito con anterioridad, sin embargo, el segundo no, por lo que el SNP-Y localizado en la posición -605 de la secuencia, es la primera vez que se determina. Así, el alelo nuevo consta de: una timina en el SNP-W (el codón pasa de GGT a GGA), y una citosina en el SNP-Y (el codón pasa de ser ATG a ACG). Al igual que el alelo nuevo 1, mencionar que, a lo largo del resto de exones, pueden existir otras variaciones nucleotídicas, a parte de las mencionadas.

- **Secuencia aminoacídica:**

Se ha determinado la secuencia aminoacídica de dicho alelo nuevo, gracias al programa MEGA 7 (Kumar et al., 2015).

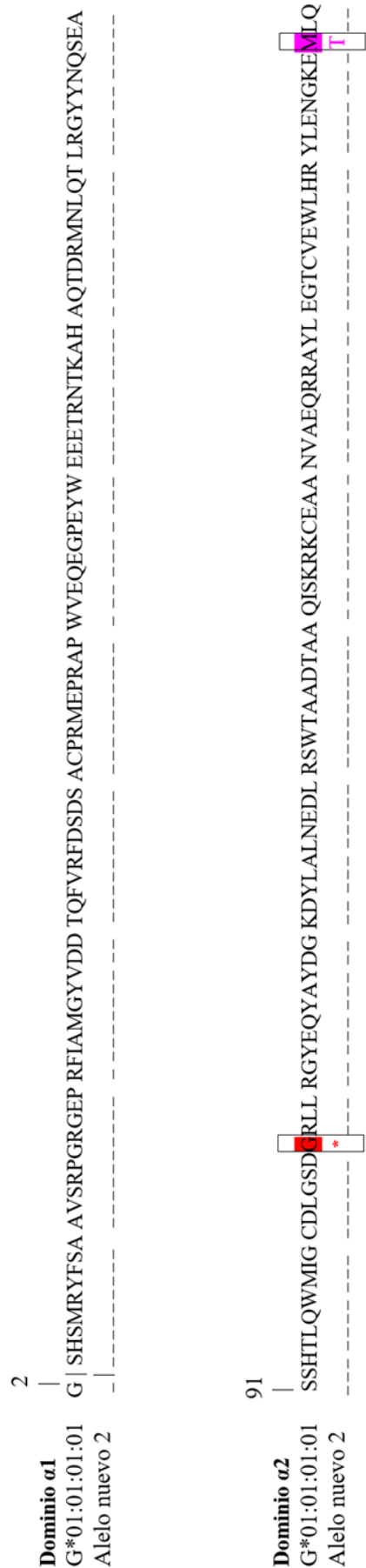


Figura R9: Secuencia aminoacídica del alelo nuevo 2 (individuo de R. Dominicana), comparada con una secuencia de referencia (HLA-G\*01:01:01), sin intrones. Se indica con guiones (-) los aminoácidos iguales a la secuencia de referencia. Las barras verticales (|) indican inicio de la secuencia. El asterisco (\*) en rojo, indica una mutación silenciosa en el triplete que codifica para ese aminoácido que coincide con la secuencia de referencia. El color rosa, indica una mutación no silenciosa en el triplete que codifica para ese aminoácido que, en este caso, no coincide con la secuencia de referencia. El número al inicio de cada línea, indica el número de posición de ese aminoácido en la secuencia. Ver correspondencia de letras en Anexos III.

Dominio $\alpha 2$		
Residuo	107	178
G*01:01:01	Gly	Met
Alelo nuevo R. Dominicana	Gly	Thr
Cambios en codones ADN	GGA→GGT	ATG→ACG

Figura R10: Variabilidad de aminoácidos entre el alelo nuevo de R. Dominicana y el alelo HLA-G\*01:01. La secuencia de este alelo fue enviada al GenBank el 29 de enero del 2015; el número de acceso KP739974 fue dado el 4 de febrero del 2015. El consenso de aa está basado en la secuencia aminoacídica de HLA-G\*01:01.

	Exon 2										Exon 3						
	13	27	31	34	54	100	104	105	107	110	159	169	171	178			
HLA-G*01:20									GGT					ACG			
HLA-G*01:01	TCC Ser	TAC Tyr	ACG Thr	GTG Val	CAG Gln	GGC Gly	GGG Gly	TCC Ser	GGA Gly	CTC Leu	TAC Tyr	CAC His	TAC Tyr	ATG Met			
HLA-G*01:02					CGG Arg												
HLA-G*01:03			TCG Ser														
HLA-G*01:04										ATC Ile							
HLA-G*01:06																	
HLA-G*01:07	TTC Phe									ATC Ile							
HLA-G*01:08																	
HLA-G*01:09											CAC His						
HLA-G*01:10			ATG Met														
HLA-G*01:11			ATG Met							ATC Ile							
HLA-G*01:12		CAC His															
HLA-G*01:14					GAC Asp												
HLA-G*01:15						GTG Val				ATC Ile							
HLA-G*01:16							TGC Cys										
HLA-G*01:17												CGC Arg	CAC His				
HLA-G*01:18																	
HLA-G*01:19				CTG Val						ATC Ile							

HLA-G\*01:20 (alelo nuevo 2 de R. Dominicana): solo se indican los codones, mutaciones sinónimas y los aa codificados.

Figura R11: Proteínas de los alelos HLA-G incluido el nuevo alelo dominicano. Se incluyen las posiciones variables de los aa de cada alelo HLA-G.

En la figura R9 se muestra la secuencia de aa nueva del individuo de origen dominicano. Para un mejor análisis de la misma, se ha comparado con una secuencia de referencia (HLA-G\*01:01:01:01), sin introducir intrones en ninguna.

De tal forma, en la figura se muestran los dominios  $\alpha 1$  y  $\alpha 2$  (codificados por el exón 2 y 3 respectivamente). En este caso, los dos cambios a nivel nucleotídico ocurren en el exón 3 (ver figura R8) por lo que ahora, es en el dominio  $\alpha 2$  donde hay que prestar atención. El primer cambio de nucleótido (SNP-W) en el residuo 107: glicina→glicina, de la secuencia aminoacídica (ver figura R9), produce una sustitución silenciosa por lo que no afecta a la proteína, como se muestra en la figura R10. El segundo cambio de nucleótido (SNP-Y), produce una sustitución no silenciosa en el residuo 178, pudiendo variar entre metionina (ATG) (como ocurre en el resto de alelos hasta ahora descritos) o treonina (ACG), en el caso de este alelo (ver figura R10). Este cambio polimórfico en la posición 178, no se ha visto anteriormente en ningún alelo descrito de HLA-G. En la figura R11 se muestran los cambios producidos en la nueva proteína (alelo nuevo 2 de R. Dominicana) comparándolos con los cambios del resto de proteínas de HLA-G.

El haplotipo extendido de este individuo es: HLA-A\* 11:01,\* 24:02; HLA-B\* 35:01,\* 35:02; HLA-DRB1\*01:01,\* 11:04; HLA-DQB1\*03:01,\*05:01; HLA-G\*01:04:01.

## **5.2 Distancias genéticas.**

Una vez determinada la secuencia nucleotídica de ambos alelos, se ha comparado la distancia genética que existe entre estas nuevas secuencias y el resto de alelos ya descritos actualmente. Para ello, se usó de nuevo el programa MEGA 7 (Kumar et al., 2015). La distancia genética entre dos secuencias de ADN se calcula en base a las diferencias que se observan entre ellas.

Los valores superiores a 0,001 no se muestran en las tablas.

**Tabla R12:** Distancias genéticas entre el alelo nuevo 1 y los alelos más cercanos a él.

Alelos	Distancias genéticas
Alelo Nuevo 1	-
G*01:01:17	0,000
G*01:01:02:01	0,001
G*01:01:02:02	0,001
G*01:01:18	0,001
G*01:01:19	0,001
G*01:06	0,001
G*01:08	0,001
G*01:18	0,001

En la tabla R12, se muestran las distintas distancias genéticas entre el alelo nuevo 1 y los ocho alelos más cercanos a él (distancias con los valores más cercanos a cero). Este alelo parece estar más próximo al grupo de alelos HLA-G\*01:01, aunque también se sugiere una cercanía genética con los grupos HLA-G\*01:06, -G\*01:08 y -G\*01:18.

Concretamente parece que el alelo nuevo, tiene más proximidad con los alelos HLA-G\*01:01:17 y -G\*01:08, con los que sólo se observa un cambio a lo largo de los tres exones estudiados. Con el resto de alelos tiene dos o tres cambios nucleotídicos.

Es necesario aclarar que, aunque los datos muestren una distancia genética, entre el alelo nuevo y HLA-G\*01:01:17, de un 0,000 existe un cambio de nucleótido entre las secuencias de ambos, esto puede ser debido al programa utilizado.

**Tabla R13:** Distancias genéticas entre el alelo nuevo 2 y los alelos más cercanos a él.

Alelos	Distancias genéticas
Alelo nuevo 2	-
G*01:01:03:01	0,001
G*01:01:03:02	0,001
G*01:01:03:03	0,001

En la tabla R13, se muestran las distintas distancias genéticas entre el alelo nuevo 2 y los tres alelos más cercanos a él (distancias con los valores más cercanos a cero). Se observa como este alelo nuevo, es más cercano al grupo de alelos HLA-G\*01:01, concretamente al HLA-G\*01:01:03, dado que el tercer par de dígitos indica mutaciones silenciosas en la región codificante que no suponen alteraciones en la secuencia aminoacídica y el cuarto par de dígitos (HLA-G\*01:01:03:**01**; -G\*01:01:03:**02**; -

G\*01:01:03:**03**), indica cambios en la región no codificante, que en este momento no ha sido analizada y que tampoco induce cambios en la secuencia de la proteína (Robinson et al., 2015).

# DISCUSIÓN





# **1. HLA-G Y SU POLIMORFISMO EN LA POBLACIÓN AMERINDIA INMIGRANTE EN MADRID.**

Una de las características principales del locus HLA-G, es su bajo polimorfismo. Se ha sugerido que esta característica podría ser debido a su función tolerogénica, así la evolución ha trabajado eliminando cambios (menor número de proteínas diferentes de HLA-G) para mantener su papel inmunológico. De tal manera, la mayoría de los alelos, se han generado por mutaciones silenciosas en exones o intrones. (Martínez-Laso et al., 2013). Sin embargo, cada vez van aumentando el número de alelos HLA-G. Existe la posibilidad de que haya un polimorfismo muy extenso también, pero que no se detecte fácilmente, pues no hay un método automático de detección de alelos, por ello, estos alelos deben ser buscados con secuenciación génica. Asimismo, cada vez se va encontrando más alelismo en el exón 4 (dominio  $\alpha 3$ ), probablemente sea debido a que las interacciones funcionales con otras moléculas de células NK y con CD8 se efectúa en este dominio (Donadi et al., 2011); estos contactos en el dominio  $\alpha 3$  de HLA-G puede ser el reflejo de su principal función y el lugar de concentración del alelismo. Es necesario mencionar otros dominios que no sean  $\alpha 1$  ni  $\alpha 2$ , ya que la valva presentadora es factible que tenga una función secundaria y sea el dominio  $\alpha 3$  y otros, los que determinen la función principal de estas moléculas. Se puede pensar, pues, que finalmente, nos encontremos con un polimorfismo elevado en la molécula HLA-G y en las otras moléculas no clásicas, HLA-E y HLA-F.

Con solo 53 alelos y 18 proteínas descritos hasta la fecha, muchos de ellos solo se han descrito en un único estudio, no observándose en más poblaciones. Estudios anteriores demuestran que poblaciones caucásicas, africanas y asiáticas, cuentan con menos variabilidad alélica dentro de este locus, que en los loci HLA de clase I clásicos, dado que son los más polimórficos del genoma humano. (Meyer & Thomson, 2001).

En el caso particular de las distintas poblaciones indígenas amerindias, todas tienen en común un limitado número de alelos HLA de clase I, clásicos (Moscoso, 2006), pero existen pocas referencias acerca de las frecuencias de alelos HLA de clase I, no clásicos, como es el locus HLA-G, en estas poblaciones (Arnaiz-Villena et al., 2013; Castelli et al., 2007).

En términos prácticos, solo cuatro proteínas HLA-G diferentes se observan normalmente en las poblaciones mundiales, codificadas por los grupos de alelos más

frecuentes; HLA-G\*01:01, G\*01:03, G\*01:04 y G\*01:06, que pueden influir en la función de HLA-G (Donadi et al., 2011).

Concretamente, en la población que este estudio comprende, se han detectado dieciocho alelos distintos no proteicos, con seis dígitos de resolución, (ver tabla R3) de los cuales el más frecuente (31,2%) es el HLA-G\*01:01:01, como ocurre en muchas de las poblaciones ya estudiadas (Brasil, China, Ghana, Polonia, Corea...) (Gonzales-Galarza et al., 2011). Dentro del grupo de alelos HLA-G\*01:01, se han detectado once alelos diferentes y dentro del grupo de alelos HLA-G\*01:04, cuatro alelos diferentes. Por lo que, quince de los dieciocho alelos encontrados, pertenecen a los grupos mundialmente más frecuentes, HLA-G\*01:01 y HLA-G\*01:04. En la tabla R2, se muestra como la frecuencia de ambos (71,8% y 18,5% respectivamente), son las más altas de la muestra. Los grupos de alelos HLA-G\*01:03 y HLA-G\*01:06, con frecuencias 5,8% y 2,9% respectivamente, son los siguientes en la tabla de frecuencias (ver tabla R2). Esto corrobora lo mencionado anteriormente.

Se ha observado como en otras poblaciones mundiales, como coreanos (Park et al., 2012), africanos de Shona (Matte et al., 2000), danses (Hviid, 2006) etc., el número de alelos HLA-G encontrados es muy reducido. Sin embargo, en amerindios inmigrantes en Madrid, el número es relativamente alto, como ocurre en brasileños (Castelli et al., 2007). Esto podría sugerir unos niveles marcados de diversidad genética, pudiendo ser debidos a la naturaleza mixta de la muestra tomada, donde hay individuos de numerosos países, centro y sudamericanos (ver tabla R1).

En relación a la composición de la muestra, en la tabla R1 se pueden observar los dieciséis países amerindios presentes en ella, tanto centroamericanos como sudamericanos. El país más representado es Ecuador, con un porcentaje del 28,8%, seguido de Colombia (26,9%), Perú (13,5%) y Bolivia (6,4%). Como se ha comentado en la Introducción, alrededor de un 4,0-4,5% de la población de la Comunidad de Madrid es amerindia (Parga-Lozano et al., 2011), y según el barómetro de inmigración de dicha comunidad, en el 2014, el país amerindio más representado en la comunidad inmigrante fue Ecuador, seguido de Colombia, Perú, Bolivia etc. Estos datos corroboran lo analizado en la muestra de estudio. Si bien es posible que haya una pequeña introducción de genes no amerindios en la población utilizada. Esta fue escogida cuidadosamente en cuanto a procedencia, de zonas rurales, con tres generaciones

estables y recién llegados de inmigración directos a Madrid. Este procedimiento se encuentra detallado en un trabajo del equipo anterior (Parga-Lozano et al., 2011).

## **2. ANÁLISIS DE LOS HAPLOTIPOS EXTENDIDOS.**

Los genes HLA se transmiten por segregación mendeliana y las variantes alélicas se expresan de modo codominante. Al grupo de alelos de loci adyacentes, presentes en cada cromosoma, se le denomina haplotipo HLA. La probabilidad de que dos hermanos tengan los mismos haplotipos HLA es del 25%, que sean diferentes otro 25% y de que compartan un solo haplotipo, 50% (Cardozo et al., 2014), si no ha ocurrido sobrecruzamiento en meiosis paterna (muy raramente ocurre en esta zona cromosómica).

Hay que mencionar que, en este caso, los haplotipos extendidos están formados por cuatro loci, HLA-A-B-DRB1-G. El último de ellos es el locus HLA-G, que como información adicional se ha añadido al haplotipo, aunque no se suele mencionar en estudios de genética de poblaciones, por ello cuando se habla de otras poblaciones no se incluye dicho locus.

En la tabla R10 se muestran los nueve haplotipos extendidos más frecuentes en la población de estudio. Respecto al bloque HLA-A-B-DRB1, los tres haplotipos más frecuentes, HLA-A\*24:02-B\*35:01-DRB1\*14:02-G\*01:04 (2,9%), HLA-A\*02:01-B\*35:04-DRB1\*08:02-G\*01:01 (1,7%) y HLA-A\*02:01-B\*40:02-DRB1\*04:07-G\*01:01 (1,2%) son característicos de amerindios, observados en poblaciones aymaras (Arnaiz-Villena et al., 2005), quechuas (Martínez-Laso et al., 2006), jaidukama (Martínez-Laso et al., 2011), mayas (Gomez-Casado et al., 2003), nahuas (Vargas-Alarcón et al., 2007), wayus (Silvera et al., 2011), lamas (Moscoso et al., 2006) y mayos (Arnaiz-Villena et al., 2007). Todos con un posible origen amerindio. Dos de ellos, portan los alelos DRB1\*08:02 y DRB1\*14:02 que se consideran verdaderos marcadores de las poblaciones (Vargas-Alarcón et al., 2000; Yunis et al., 1994) también, el alelo DRB1\*04:07, portado con frecuencia por amerindios (Moscoso, 2006). Esto apoya la determinación del origen amerindio de los individuos analizados para el estudio. Asimismo, el haplotipo HLA-A\*24:02-B\*35:01-DRB1\*14:02 también ha sido observado por Parga-Lozano y colaboradores en la población amerindia inmigrante, como uno de los haplotipos extendidos más frecuentes de esa población. (Parga-Lozano et al., 2011).

El perfil HLA amerindio es, en conjunto, diferente al resto de poblaciones mundiales, asimismo existen alelos quasi específicos para amerindios. Esto también se ha visto en la población inmigrante en Madrid (Parga-Lozano et al., 2011). Sin embargo, una entrada genética europea se puede apreciar en el haplotipo HLA-A\*03:01-B\*07:02-DRB1\*15:01-G\*01:01, con posible origen de Europa occidental, que ha sido observado, concretamente, en poblaciones del norte de España, con frecuencias elevadas (Sánchez-Velasco et al., 2003).

Los haplotipos HLA-A\*02:01-B\*44:03-DRB1\*13:02-G\*01:01, HLA-A\*03:01-B\*51:01-DRB1\*04:04-G\*01:01, HLA-A\*24:02-B\*14:02-DRB1\*07:01-G\*01:01 y HLA-A\*68:01-B\*40:04-DRB1\*04:03-G\*01:01 (ver tabla R10), sin embargo, no se han encontrado en ninguna región o continente con frecuencias lo suficientemente altas como para determinar un posible origen. A pesar de ello, se han estudiado los loci HLA-A (como representación de la clase I) y HLA-DRB1 (como representación de la clase II) para poder obtener más información. Respecto al locus HLA-A, se ha demostrado que todos los alelos HLA de clase I de las poblaciones amerindias, podrían ser originadas de cuatro alelos “fundadores” HLA-A: -A\*02:01, -A\*24:02, -A\*31:01:02 y -A\*68:01:02 (Moscoso, 2006). El alelo -A\*02:01 forma parte de tres de los haplotipos más frecuentes de la población de estudio, uno de ellos en el haplotipo HLA-A\*02:01-B\*44:03-DRB1\*13:02-G\*01:01, por lo que, a pesar de no tener un origen claro, si parece estar relacionado con la población amerindia. Lo mismo ocurre con los alelos A\*24:02 y A\*68:01 presentes en los haplotipos, HLA-A\*24:02-B\*14:02-DRB1\*07:01-G\*01:01 y HLA-A\*68:01-B\*40:04-DRB1\*04:03-G\*01:01, respectivamente. Acerca del locus HLA-DRB1, uno de los loci más usado para el estudio de la genética de poblaciones a nivel mundial, es sabido que los grupos mestizos (mezcla de amerindios con europeos) de América, han sido menos estudiados en comparación con las poblaciones nativas. En esta comparación, se observó que los grupos serológicos HLA eran compartidos por los dos grupos poblacionales, pero no compartían los mismos alelos. Los amerindios portaban los alelos: HLA-DRB1\*04:07, DRB1\*04:03 y DRB1\*04:11, mientras que los mestizos portaban los siguientes: DRB1\*14:01, DRB1\*07:01, DRB1\*04:04 y DRB1\*08:01, como los más característicos de estas poblaciones (Petzl-Erler et al., 1997). El haplotipo HLA-A\*68:01-B\*40:04-DRB1\*04:03-G\*01:01, posee el alelo DRB1\*04:03 propio de amerindios, por lo que ya es un haplotipo asociado a este grupo poblacional (Petzl-Erler et al., 1997), y los

haplotipos HLA-A\*03:01-B\*51:01-DRB1\*04:04-G\*01:01 y HLA-A\*24:02-B\*14:02-DRB1\*07:01-G\*01:01, portan los alelos DRB1\*04:04 y DRB1\*07:01, respectivamente, característicos del grupo mestizo, lo que podría indicar que la muestra de estudio, efectivamente, ha sufrido un intercambio genético con poblaciones europeas.

Por último, se han detectado cuatro haplotipos nuevos en la población de estudio, que no han sido descritos en ninguna otra población. Estos haplotipos son: HLA-A\*31:01-B\*39:06-DRB1\*13:02-G\*01:01; HLA-A\*29:02-B\*15:05-DRB1\*04:03-G\*01:05N; HLA-A\*31:01-B\*15:05-DRB1\*04:03-G\*01:05N y HLA-A\*33:03-B\*57:03-DRB1\*13:02-G\*01:05N (ver tabla R10). Es obvio decir que, todos los haplotipos extendidos de este trabajo, en los que se han incluido alelos HLA-G, son nuevos (tabla R10).

El locus HLA-G como parte de los haplotipos extendidos en amerindios es una información añadida muy valiosa. Los genes de histocompatibilidad de amerindios son radicalmente diferentes a los del resto de poblaciones mundiales. Los genes HLA clásicos y no clásicos, intervienen en el trasplante de órganos y en el ligamiento de HLA a enfermedad. Por ello, es necesario determinar ambos en amerindios, sobre todo los genes HLA no clásicos, ya que se han investigado en menor medida, en particular HLA-G. Así, toda información, ya sea del polimorfismo alélico, como de haplotipos extendidos en esta población, será muy útil para poder crear una lista de espera de trasplante virtual. También se podrán establecer programas epidemiológicos específicos para HLA asociado a enfermedad.

### **3. ANÁLISIS DEL ALELO NULO, HLA-G\*01:05N, EN LA POBLACIÓN AMERINDIA.**

Como se muestra en los resultados (ver apartado 4, página 106), se han detectado tres alelos nulos en tres individuos diferentes, todos ellos en heterocigosis. Proceden de Colombia, Argentina y Nicaragua, dos hombres y una mujer, respectivamente.

En 1987, Suárez y colaboradores, en nuestro mismo grupo de investigación, describieron el primer alelo nulo del locus HLA-G en la población española (Suarez et al., 1997). Este alelo nulo, HLA-G\*01:05N, presenta la delección de una citosina en el último nucleótido del codón 129 o primer nucleótido del codón 130, del exón 3, causando un cambio en la pauta de lectura con la aparición de un codón de terminación en ambos codones, codón 189 (exón 4), que bloquea la traducción de HLA-G1 y -G5, o

en el codón 297 en el exón 5, bloqueando la traducción de HLA-G4. Sin embargo, este alelo nulo es capaz de mantener la traducción de las isoformas ancladas a membrana HLA-G2 y -G3 y las solubles HLA-G6 y -G7, ya que en todas ellas el exón 3, que contiene la delección, se ha eliminado por splicing alternativo. (Le Discorde et al., 2005; Suarez et al., 1997).

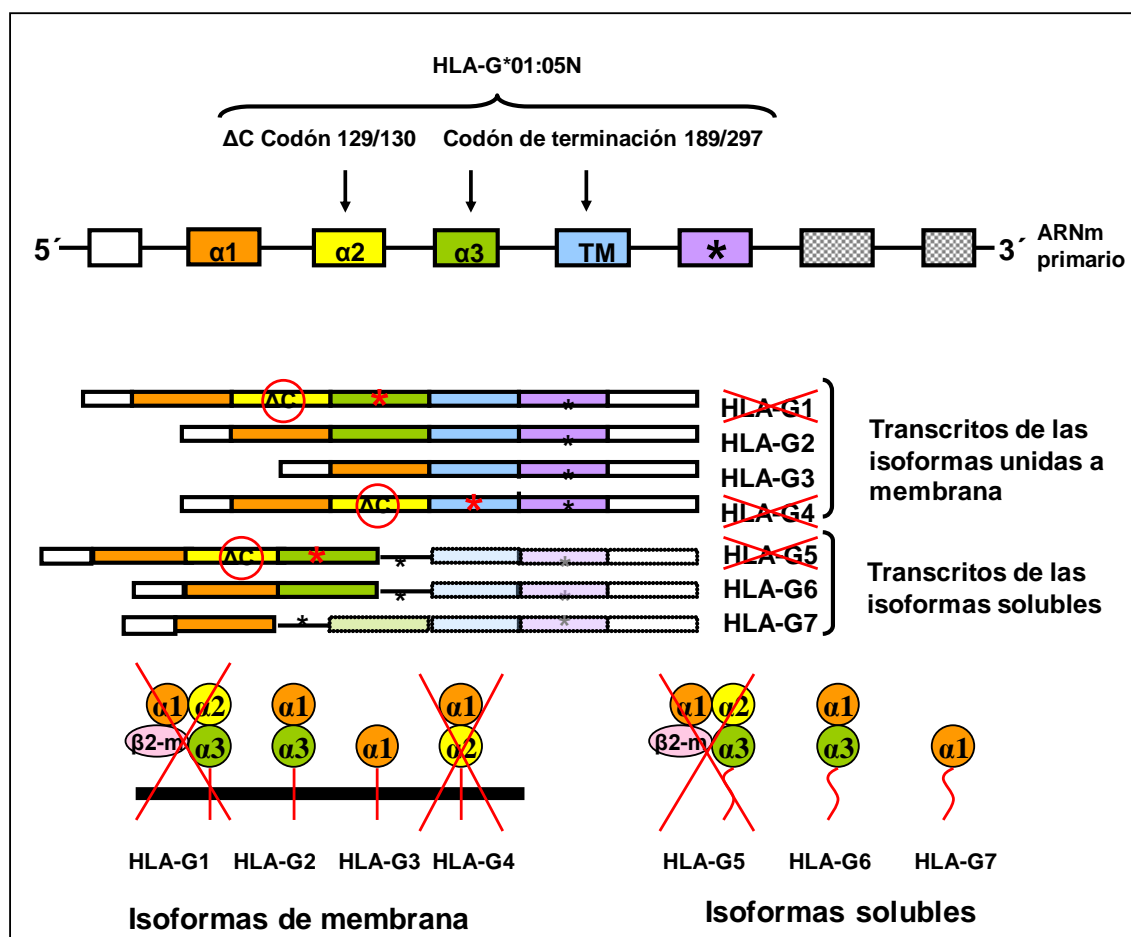


Figura D1: Figura el gen de HLA-G, correspondiendo al dominio del péptido líder (exón 1), dominios α1, α2 y α3 (exones 2, 3 y 4), dominio transmembrana (exón 5) y el dominio citoplasmático (exón 6 con el codón de terminación prematuro). Las 7 isoformas de HLA-G producidas por splicing alternativo del ARNm primario. El gen HLA-G\*01:05N codifica para las isoformas de membrana HLA-G2, -G3 y solubles HLA-G6 y -G7. (Donadi et al., 2011)

### 3.1 Alelo nulo en Uros y Mayas.

Los uros y los mayas pertenecen a poblaciones amerindias aisladas, concretamente el grupo étnico de los uros proviene de las islas flotantes del Lago Titikaka (Altiplano de Perú) y el grupo de los mayas, al pueblo de Quetzaltenango (Guatemala).

El análisis de la frecuencia de este alelo en dichas poblaciones, se ha realizado en este equipo de trabajo (Arnaiz-Villena et al., 2013).

El alelo nulo HLA-G\*01:05N fue descrito en españoles en 1997 (Suarez et al., 1997) y ha sido asociado con abortos espontáneos y recurrentes. Por otro lado, también ha sido postulado el efecto protector contra infecciones intra-uterinas durante la gestación, por lo que, siendo un tema aun sin esclarecer, ambas circunstancias podrían coexistir (Arnaiz-Villena et al., 2013).

Lo más llamativo es la ausencia absoluta del alelo HLA-G\*01:05N tanto en uros como en mayas. Esto contradice la teoría de que las frecuentes infecciones intra-uterinas pueden mantener las frecuencias del alelo nulo, dado que la atención sanitaria en las comunidades amerindias es (y ha sido) deficiente, comparativamente a los europeos. Sin embargo, estos resultados son concordantes con otras poblaciones amerindias, cuya pureza está más en duda, pero presenta este alelo en baja frecuencia, por ejemplo la población brasileña general urbana (Castelli et al., 2007).

### **3.2 Alelo nulo en amerindios inmigrantes en Madrid.**

La frecuencia de este alelo en la muestra de estudio, con un 1,0%, no ha sido muy elevada (ver tablas R2). Es similar a lo que se observa en otras poblaciones amerindias como uros y mayas, donde la frecuencia es inexistente (ver tabla R11). Por el contrario, las poblaciones con una frecuencia significativamente mayor, son algunas caucásicas de Oriente Medio (iraquíes, iraníes, indios del Norte de India) y ciertas poblaciones africanas (Ghana, Shona y afroamericanos) (ver tabla R11 y figura R3).

La secuencia del alelo HLA-G\*01:05N podría indicar que este ha sido originado a partir del alelo HLA-G\*01:01:02, al producirse la delección de la cisteína en el codón 129/130, ya que sus secuencias son idénticas, excepto dicha delección (Castro et al., 2000). Adicionalmente, HLA-G\*01:05N está asociado al haplotipo HLA-A\*30:01-B\*13:02 frecuente en Oriente Medio y algunas poblaciones mediterráneas, ya que pudo haber sido introducido en España por invasores árabes en el siglo VIII DC (Suarez et al., 1997), lo que podría situar a Oriente Medio como lugar de origen de esta mutación (Arnaiz-Villena et al., 2013).

Respecto a los haplotipos extendidos estudiados, tres, de toda la lista de haplotipos obtenidos, portaban el alelo nulo; HLA-A\*33:03-B\*57:03-DRB1\*13:02-G\*01:05N, HLA-A\*31:01-B\*15:05-DRB1\*04:03-G\*01:05N y HLA-A\*29:02-B\*15:05-DRB1\*04:03-G\*01:05N. Todos ellos, con baja frecuencia (ver tabla R10), lo que sigue mostrando la baja presencia de este alelo en poblaciones amerindias. Estos tres haplotipos no se han encontrado en ninguna otra población (ver tabla R10), por lo que



se podría sugerir que son haplotipos nuevos, sin embargo, teniendo en cuenta, principalmente la baja frecuencia de estos, hay que tomar con cautela estos datos ya que, como se ha mencionado anteriormente, podrían ser un mero artefacto estadístico, no existiendo en la biología humana.

Por otro lado, existen algunas combinaciones de genes en los haplotipos, que se presentan con mayor frecuencia de lo esperado, basándonos solamente en las frecuencias génicas de la población. Este fenómeno recibe el nombre de desequilibrio de ligamento y dentro del sistema HLA desaparece muy despacio; podría ser usado para reconocer singularizar poblaciones (Degos & Dausset, 1974). En concreto, el desequilibrio de ligamiento entre el locus..... HLA-G y HLA-A es bien conocido y fue descrito ya en 1993 (Morales et al., 1993). Esto ocurre debido a la distancia física tan corta que separa a los dos loci dentro del cromosoma, que parece no ser superior a las 300kb (Morales et al., 1993). Por lo que sugiere, que el proceso de recombinación entre ambos es poco frecuente, manteniendo a lo largo de la evolución las combinaciones de alelos HLA-A y -G.(Ober et al., 1996)

Respecto a los datos obtenidos en este estudio, el alelo nulo estaría en desequilibrio de ligamiento con HLA-A\*30:01 y -A\*31:01 (ver tabla R5) y con HLA-B\*15:05 y -B\*57:03 (ver tabla R7). Se vuelve a observar como HLA-G\*01:05N estaría asociado al alelo HLA-A\*30:01 (Suarez et al., 1997) que presenta frecuencias elevadas en regiones orientales, africanas y caucasoides (Gonzales-Galarza et al., 2011), apoyando así, la teoría de un origen mediterráneo oriental. En el caso del desequilibrio de ligamiento entre el locus HLA-G y HLA-B, la distancia física entre ambos es mayor que la existente entre HLA-G y -A, ya que HLA-B se sitúa más en el extremo centromérico y HLA-G más en el extremo telomérico del cromosoma dentro de la región de clase I. Por ello, los datos pueden ser un tanto más difusos. El alelo HLA-B\*15:05, tiene un posible origen amerindio (Gonzalez-Galarza et al., 2011), y dos de los haplotipos extendidos llevan este alelo, HLA-A\*31:01-B\*15:05-DRB1\*04:03-G\*01:05N y HLA-A\*29:02-B\*15:05-DRB1\*04:03-G\*01:05N (tabla R10). El alelo -B\*57:03 tiene un posible origen negro, caucasioide (Gonzalez-Galarza et al., 2011) y se observa en el haplotipo extendido HLA-A\*33:03-B\*57:03-DRB1\*13:02-G\*01:05N. El desequilibrio de ligamiento puede ser, principalmente, un indicativo de que ha tenido lugar una fusión de poblaciones en esa región, (Degos & Dausset, 1974), por lo que se podría entender que



ha existido un mestizaje en la población de estudio, probablemente con población europea y negroide de África.

La existencia de genes “enfermos” en una población suele ser justificada por ventajas evolutivas. Como se ha comentado anteriormente, la molécula inmunosupresora HLA-G, juega un papel primordial en la aceptación del feto por la madre durante la gestación. Los intereses materno-fetales pueden llegar a ser contradictorios cuanto menores sean los niveles de HLA-G, el sistema inmunológico de la madre podría provocar daños en el feto y finalmente abortos espontáneos en madres heterocigotas u homocigotas para el alelo HLA-G\*01:05N (Arnaiz-Villena et al., 2013). Sin embargo, Castro y colaboradores observaron cómo mujeres de una misma familia española, homocigotas para el gen HLA-G\*01:05N, eran personas sanas capaces de tener partos totalmente normales, indicando que la isoforma HLA-G1 no es imprescindible para la supervivencia humana (Castro et al., 2000). Como se ha mencionado, se han descrito efectos protectores de este alelo, frente a infecciones intra-uterinas en mujeres embarazadas. No está claro, pero ambas circunstancias podrían coexistir (Arnaiz-Villena et al., 2013). La presencia de este alelo nulo en la población, podría ser debido a que el dominio  $\alpha 1$  es suficiente para el normal funcionamiento de la molécula HLA-G, de tal forma las fuerzas evolutivas negativas no actuarían eliminando este gen.

En primates, todos los individuos estudiados del género *Macaca* tampoco tiene completa la molécula de MHC-G (Arnaiz-Villena et al., 1999), aunque la mutación producida es distinta a la ocurrida en la especie humana, en ambos casos se pueden expresar los dominios  $\alpha 1$ , tanto soluble como de membrana. Esto indicaría que el alelo nulo confiere alguna ventaja en la población (Arnaiz-Villena et al., 2013). Se ha sugerido la posibilidad de una ventaja evolutiva, tanto en la madre como en el feto, dado que estos monos también tienen este alelo nulo (Aldrich et al., 2002). Como se ha mencionado con anterioridad, el alelo HLA-G\*01:05N podría proporcionar cierta ventaja selectiva, no sólo a nivel materno-fetal, sino al mejorar la respuesta inmune frente a la infección por VIH (Lajoie et al., 2006) u otras infecciones del sistema inmune no relacionadas directamente con la gestación.

Estos resultados pueden ser importantes en el campo de la epidemiología de infecciones en amerindios, si se llevan a cabo estudios en patología fetal y abortos

espontáneos, así como en medicina preventiva en la relación materno-fetal, tanto sana como patológica.

#### **4. ALELOS NUEVOS Y SUS POSIBLES IMPLICACIONES A NIVEL MOLECULAR.**

Como se muestra en los resultados (ver apartado 5), en este estudio se han descubierto dos alelos nuevos, ambos en individuos diferentes. El primero de ellos, de origen ecuatoriano y el segundo, de origen dominicano.

En relación al alelo nuevo 1 de Ecuador, los cambios descritos en la secuencia aminoacídica se encuentran en las posiciones 57, 93, 100 y 219 (ver figura R5). Las tres primeras posiciones son producto de mutaciones silenciosas en la secuencia de ADN, por lo que el aa no cambia en ninguna de ellas. Aunque hay que mencionar que es el polimorfismo en el residuo 100 el que hace de este alelo, un alelo nuevo no descrito anteriormente, ya que, si no sería idéntico al alelo HLA-G\*01:08.

Por otro lado, es en la cuarta posición donde si se observa un cambio de aa, una arginina (R) (CGG) por un triptófano (W) (TGG) (ver figura R5). Este cambio también se ha observado en el alelo HLA-G\*01:08. El residuo 219, está localizada dentro del dominio  $\alpha 3$  (codificado por el exón 4) de la molécula HLA-G. Este dominio interacciona con CD8 $\alpha\alpha$  (correceptor de las células T existente en la superficie celular como homodímero con dos cadenas  $\alpha$ , CD8 $\alpha\alpha$ , y como heterodímero con una cadena  $\alpha$  y otra  $\beta$ , CD8 $\alpha\beta$ ) a través de una región complementaria formada por un lazo localizado entre los residuos 223 y 229 del dominio (Gao et al., 2000). Los polimorfismos en esta región o zonas circundantes, como es el caso del residuo 219, podrían afectar a la interacción entre el CD8 $\alpha\alpha$  y el dominio  $\alpha 3$  de la molécula HLA-G (Gao et al., 2000). Aunque se sabe que la afinidad de HLA-G por CD8 $\alpha\alpha$  es comparable a la existente entre HLA-A\*02/CD8 $\alpha\alpha$ , en contraste a la baja afinidad con HLA-E (Gao et al., 2000), la afinidad de esta interacción está muy relacionada con el tipo de alelo considerado. Por lo que, a pesar de que la afinidad en muchos de los alelos HLA-G, no ha sido estudiada, aquellos que exhiben diferentes residuos en el dominio  $\alpha 3$  podrían presentar diferencias en la afinidad al correceptor CD8 $\alpha\alpha$  (Donadi et al., 2011).

Por otro lado, hay que mencionar que este dominio también es el sitio de interacción de HLA-G con los receptores LILRBs, expresados en un amplio rango de células del sistema inmunológico, incluyendo las células NK y linfocitos T. De tal forma, se sabe

que el residuo 219 tiene una función en la avidez de unión de estos receptores a los distintos alelos HLA de clase I. Además, los receptores LILRBs compiten directamente con CD8 $\alpha$  para unirse a las moléculas HLA de clase I a través del dominio  $\alpha$ 3 (Shiroishi et al., 2003).

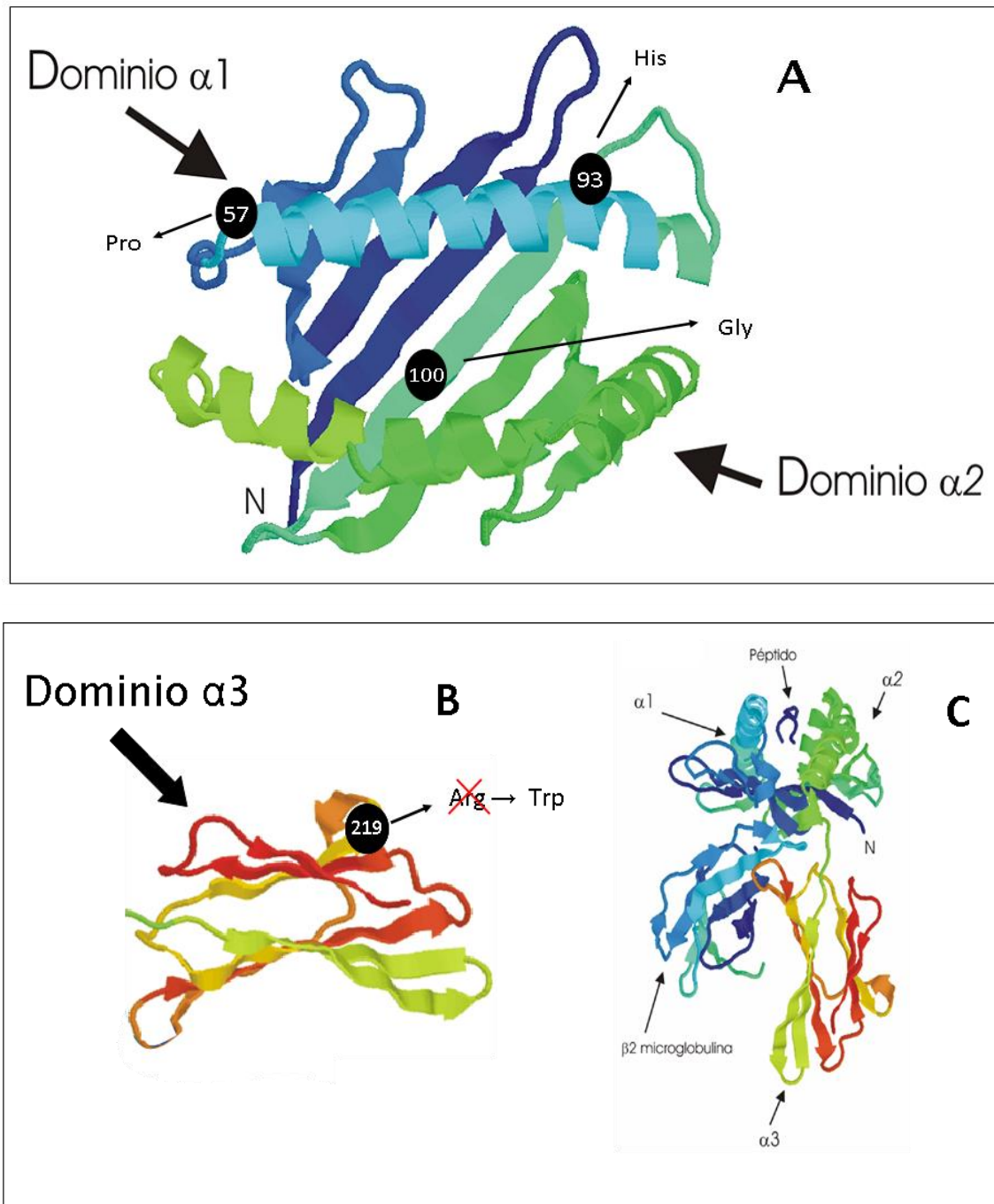


Figura D2: Estructura terciaria de la molécula HLA de clase I. Localización de los residuos 57, 93, 100 y 219 del alelo nuevo 1 en la molécula. A- El residuo 57 se encuentra en el dominio  $\alpha$ 1, los residuos 93 y 100 en el dominio  $\alpha$ 2. B- El residuo 219 se encuentra en el dominio  $\alpha$ 3, que no forma parte de la valva de unión de péptidos. C- Estructura completa con un péptido asociado y la  $\beta$ 2-m.(Gomez-Casado et al., 2006)

Por lo tanto, todo ello podría sugerir que este nuevo alelo codificase para una proteína nueva que podría tener una afinidad distinta por el correceptor CD8 $\alpha\alpha$ , pudiendo alterar, por ejemplo, su papel en la protección al feto frente al rechazo de la madre mediante la supresión de linfocitos T CD8+. Por otro lado, también podría verse afectada la interacción entre la molécula HLA-G codificada por este nuevo alelo y los receptores LILRBs, pudiendo modificar así, su implicación en distintos campos como el control de células tumorales o en rechazo del feto en la gestación.

En relación al origen de este nuevo alelo, es muy probable que surgiera de los primeros habitantes amerindios. Sin embargo, no se puede descartar que se pudiese haber dado en la población europea, dada la mezcla que tuvieron estos últimos con el pueblo amerindio después de 1492. A pesar de ello, esta teoría no es muy probable, ya que, muchas poblaciones de europeos han sido estudiados para HLA-G. Tampoco se puede descartar, la posibilidad de que haya surgido en la población negra africana, después llevados a América como esclavos, o incluso de la población oriental. Estas dos últimas hipótesis tienen una menor verosimilitud, dado que, existen extensos estudios de polimorfismos de HLA-G en ambos tipos de poblaciones sin haber aparecido este nuevo alelo.

Respecto al alelo nuevo 2 de República Dominicana, los cambios descritos en la secuencia aminoacídica se encuentran en las posiciones 107 y 178 (ver figura R9). De estos dos, es el residuo 178 el producido por una mutación no silenciosa en la secuencia nucleotídica. Mientras que en el alelo nuevo 1, todos los cambios en la secuencia estaban determinados con anterioridad, en este alelo, el SNP-Y en la posición -605 de la secuencia nucleotídica (ver figura R4) no se había observado anteriormente. Esta mutación produce un cambio de aa, una metionina (M) (ATG) por una treonina (T) (ACG) (ver figuras R9 y R10) en la posición 178 del dominio  $\alpha 2$ .

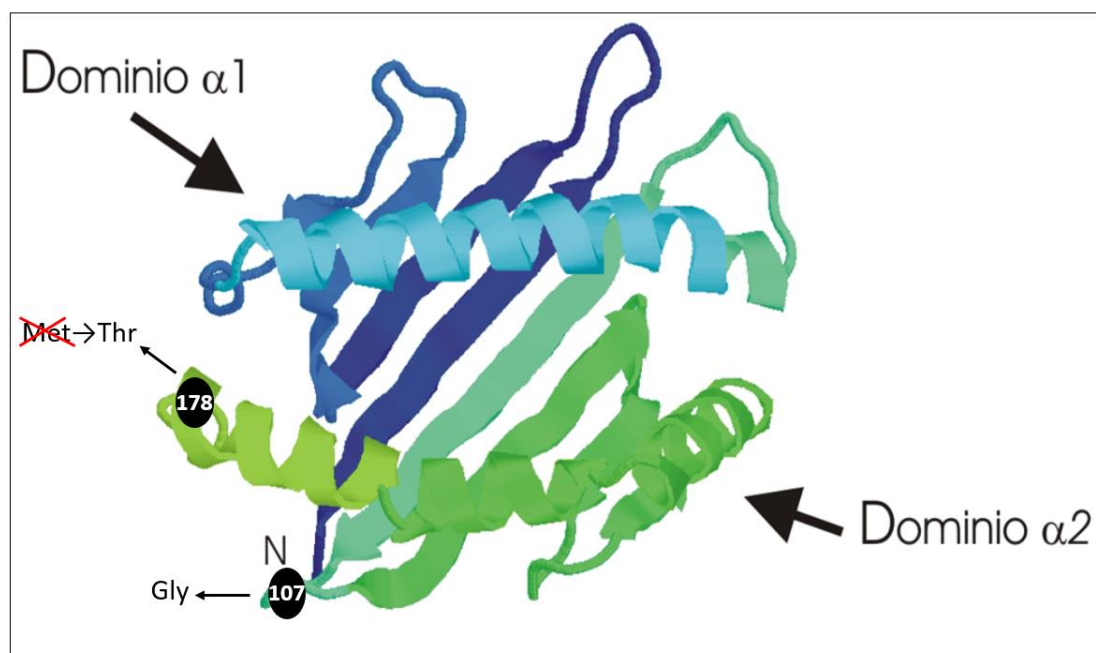


Figura D3: Estructura terciaria de la valva de la molécula HLA de clase I. Localización de los residuos 107 y 178 del alelo nuevo 2 en la molécula. Todos ellos se localizan en el dominio  $\alpha 2$ . (Gomez-Casado et al., 2006)

De tal manera, el residuo en posición 178 es el único que podría provocar modificaciones en la estructura terciaria de la molécula. Este residuo se encuentra dentro del dominio  $\alpha 2$  (codificado por el exón 3) que interacciona con el péptido y con el TCR, reconociendo residuos polimórficos (Tysoe-Calnon et al., 1991). El cambio de aa podría provocar ciertas modificaciones en la estructura terciaria de la molécula, dado que las características químicas de ambos difieren; las cadenas laterales de los aa hidrofóbicos, no polares, tienden a orientarse hacia el interior de la molécula evitando las interacciones con el disolvente, como es el caso de la metionina, mientras que las cadenas laterales de los aa hidrofílicos, polares, tienden a localizarse en la superficie de la molécula, interaccionando con el agua y permitiendo que la proteína permanezca soluble en este medio, como es el caso de la treonina. A pesar de ello, el residuo 178 no parece tener una posición relevante en la interacción con el péptido o el TCR en el modelo de la molécula HLA-A2. (Bjorkman et al., 1987b), por lo que no habría modificaciones en la funcionalidad de la molécula HLA-G. Sin embargo, por analogía, en la unión al TCR, el cambio está muy esquinado en la molécula y el residuo, con un cambio tan radical, podría influir en la unión al TCR con los residuos vecinos 173 y 174.

Respecto al origen de este alelo nuevo, este individuo tiene un fenotipo mixto, es decir, con características negroides, europeas caucásicas y amerindias. De hecho, un

gran porcentaje de los genes HLA amerindios, han sido encontrados recientemente en islas caribeñas, a pesar de la falta de características del fenotipo amerindio (Alegre et al., 2007).

## **5. LA FISIOLOGÍA Y PATOLOGÍA DE HLA-G.**

La molécula HLA-G se expresa específicamente en el citotrofoblasto placentario. Su función principal se ha asociado a las señales negativas o supresoras que el feto envía al sistema inmunológico de la madre (Donadi et al., 2011). Posteriormente se ha comprobado que esta molécula también: 1- puede actuar como inmunorregulador en procesos de autoinmunidad; 2- se expresa abundantemente en trasplantes de órganos, que suelen provocar poco rechazo, por ejemplo, hígado y córnea; 3- se asocia a patologías relacionadas con abortos y enfermedades del embarazo; 4- los tumores pueden defenderse del control del sistema inmune con expresión de HLA-G en células tumorales; 5- asimismo, puede intervenir en resistencia a enfermedades infecciosas (Donadi et al., 2011).

Por ello, el estudio poblacional de los alelos HLA-G puede ser útil para establecer una medicina preventiva y estudio epidemiológico de las patologías mencionadas, incluido los programas de trasplante.

El estudio de alelos HLA de otros loci (que en el presente trabajo también se ha efectuado) es útil para la asociación de HLA y farmacogenómica, HLA y ligamiento a enfermedad y HLA y trasplante (Arnaiz-Villena et al; 2017).

Sin embargo, prácticamente todos los estudios de definición de alelismo HLA-G, se han hecho en poblaciones no amerindias, que se sabe que son totalmente diferentes al resto del mundo en cuanto a perfil cualitativo (diferentes alelos) y cuantitativo (diferentes frecuencias) de HLA. Este es el primer estudio en profundidad, no solo del alelismo HLA-G, sino en establecer los haplotipos extendidos HLA-A-B-DRB1-G. Estos haplotipos extendidos HLA, se ha postulado que actúan conjuntamente para el funcionalismo HLA (Gomez-Casado et al., 2006).

En conclusión, los nuevos alelos HLA-G y haplotipos HLA-A-B-DRB1-G, aquí descritos, pudiesen ser de utilidad para los futuros programas de trasplantes de órganos (lista de espera virtual) en amerindios, y además establecer, con los mismo, unas bases de epidemiología/medicina preventiva asociada a los loci, HLA-G como a los HLA-DRB1, HLA-A y HLA-B. La farmacogenómica HLA (efectos adversos de

medicamentos asociados a genes HLA) está sin estudiar en amerindios. Es decir, todo el polimorfismo, fisiología y patología del locus HLA-G estaba sin base genética y sin estudio, ya que, las investigaciones se habían hecho en europeos, asiáticos y africanos (Arnaiz-Villena et al; 2017)

En el presente trabajo se ha comenzado a paliar esta deficiencia de estudios en amerindios.

# CONCLUSIONES





- 1- Se han descrito, por vez primera, nuevos haplotipos extendidos HLA-G/HLA-clase I/HLA-clase II en amerindios.
- 2- Se ha observado que, en contra de lo esperado, las frecuencias de alelos HLA-G en amerindios se asemejan más a las europeas que a las asiáticas.
- 3- Se registran unas frecuencias altas, de alelo nulo HLA-G\*01:05N, en poblaciones de Oriente Medio, mientras que son muy bajas en amerindios, se discuten los factores ambientales que han podido mantener esa gran diferencia y sus implicaciones funcionales y evolutivas.
- 4- Se ha encontrado un alelo nuevo HLA-G en una población ecuatoriana del Altiplano: HLA-G\*01:08:02. Existen varios cambios en la molécula: en el dominio  $\alpha 3$  (residuo 219, Arg→Trp), podría cambiar la interacción de la molécula nueva con CD8 y los receptores LILRBs, es decir, de células presentadoras con linfocitos B, T, NK, macrófagos, monocitos y células dendríticas.
- 5- Se ha encontrado un alelo nuevo HLA-G en una población del Mar Caribe (República Dominicana), HLA-G\*01:20, con un cambio importante en el dominio  $\alpha 2$  (residuo 178, Met→Thr) en una esquina de los bucles de la valva de HLA-G que puede afectar a la interacción con el receptor clonotípico de los linfocitos.
- 6- Se han observado alelos nulos HLA-G\*01:05N en baja frecuencia en amerindios de Centroamérica, Norte y Sur de Sudamérica.
- 7- Se han descrito doce nuevos haplotipos HLA-A-B-DRB1-G en amerindios. Asimismo, se han encontrado cuatro nuevos haplotipos extendidos, sin HLA-G, en la población amerindia inmigrante en Madrid.
- 8- Las nuevas moléculas y los nuevos haplotipos que se han descrito, pueden tener una relevancia en la fisiología y patología asociada a HLA y a HLA-G en amerindios.

# BIBLIOGRAFÍA



Abbas A, Tripathi P, Naik S, Agrawal S (2004) Analysis of human leukocyte antigen (HLA)-G polymorphism in normal women and in women with recurrent spontaneous abortions. *Eur. J Immunogenet.* 31:275-278.

Aimel JL (1967) Study of the leukocyte phenotypes in Hodking's disease. In: Curtoni ES, Mattiuz PL, Tosi RM eds. *Histocompatibility Testing 1967*. Copenhagen: Munksgaard, pp. 79-81.

Aldrich C, Wambebe C, Odama L, Di Rienzo A, Ober C (2002) Linkage disequilibrium and age estimates of a deletion polymorphism (1597DeltaC) in HLA-G suggest non-neutral evolution. *Hum Immunol* 63 (5): 405-412.

Alegre R, Moscoso J, Martinez-Laso J, Martin-Villa M, Suarez J, Moreno A, Serrano-Vela JL, Vargas-Alarcon G, Pacheco R, Arnaiz-Villena A (2007) HLA genes in Cubans and the detection of Amerindian alleles. *Mol Immunol* 44 (9): 2426-2435.

Alvarez M, Santos P, Martinho A, Simoes O, Abade A, Breda-Coimbra H (1999) HLA-G genetic polymorphism in 57 portuguese white families studied by PCR-RFLP and PCR-SSOP. *Transplant. Proc.* 31(4): 1829-1831.

Amodio G, Gregori S (2012). Distinctive Immunological Functions of HLA-G, Histocompatibility, Dr. Bahaa Abdel-Salam (Ed.), InTech, DOI: 10.5772/35248.

Arnaiz-Villena A, Palacio-Grüber J, Muñiz E, Campos C, Alonso Rubio J, Gomez Casado E, Salih SF, Martin-Villa JM, Al-Qadi R (2017) Genetic HLA study of Kurds in Iraq, Iran and Tbilisi (Caucasus, Georgia): Relatedness and Medical Implications. *Plos One* doi: 10.1371/journal.pone.0169929.

Arnaiz-Villena A, Enriquez-de-Salamanca M, Palacio-Grüber J, Campos C, Camacho A, Martin-Villa JM, Martinez-Quiles N, Gomez Casado E, Muñiz E. (2016) Characterisation and functional implications of the two new HLA-G alleles found in Amerindian and Caribbean populations. *Human Immunology* 77(9):812-816.

Arnaiz-Villena A, Areces C, Enríquez-de-Salamanca M, Abd-El-Fatah-Khalil S, Marco J, Muñiz E, Fernandez.Honrado M, Martin-Villa M, Rey D (2014) Pacific Islanders and Amerindian relatedness according to HLA autosomal genes. *International Journal of Modern Anthropology* 7: 44-67.

Arnaiz-Villena A, Regueiro JR, Lopez-Larrea C (1995) *Inmunología*. España: Editorial Complutense.

Arnaiz-Villena A, Enriquez-de-Salamanca M, Areces C, Alonso-Rubio J, Abd-El-Fatah-Khalil S, Fernandez-Honrado M, Rey D (2013) HLA-G(\*)01:05N null allele in Mayans (Guatemala) and Uros (Titikaka Lake, Peru): evolution and population genetics. *Hum Immunol* 74 (4): 478-482.

Arnaiz-Villena A, Martinez-Laso J, Moscoso J, Livshits G, Zamora J, Gomez-Casado E, Silvera-Redondo C, Melvin K, Crawford MH (2003) HLA genes in the Chuvashian population from European Russia: admixture of Central European and Mediterranean populations. *Hum Biol* 75 (3): 375-392.

Arnaiz-Villena A, Morales P, Gomez-Casado E, Castro MJ, Varela P, Rojo-Amigo R, Martinez-Laso J (1999) Evolution of MHC-G in primates: a different kind of molecule for each group of species. *J Reprod Immunol* 43 (2): 111-125.

Arnaiz-Villena A, Moscoso J, Granados J, Serrano-Vela JI, de la Peña A, Reguera R, Ferri A, Seclen E, Izaguirre R, Perez-Hernandez N et al (2007) HLA Genes in Mayos Population from Northeast Mexico. *Curr Genomics* 8 (7): 466-475.

Arnaiz-Villena A, Parga-Lozano C, Moreno E, Areces C, Rey D, Gomez-Prieto P (2010) The Origin of Amerindians and the Peopling of the Americas According to HLA Genes: Admixture with Asian and Pacific People. *Curr Genomics* 11 (2): 103-114.

Arnaiz-Villena A, Siles N, Moscoso J, Zamora J, Serrano-Vela JI, Gomez-Casado E, Castro MJ, Martinez-Laso J (2005) Origin of Aymaras from Bolivia and their relationship with other Amerindians according to HLA genes. *Tissue Antigens* 65 (4): 379-390.

Arnaiz-Villena A, Vargas-Alarcon G, Granados J, Gomez-Casado E, Longas J, Gonzales-Hevilla M, Zuniga J, Salgado N, Hernandez-Pacheco G, Guillen J et al (2000) HLA genes in Mexican Mazatecans, the peopling of the Americas and the uniqueness of Amerindians. *Tissue Antigens* 56 (5): 405-416.

Barber LD, Parham P (1993) Peptide binding to major histocompatibility complex molecules. *Annu Rev Cell Biol* 9: 163-206.

Bjorkman PJ, Saper MA, Samraoui B, Bennett WS (1987a) Structure of the human class I histocompatibility antigen, HLA-A 2. *Nature* 329: 506-512.

Bjorkman PJ, Saper MA, Samraoui B, Bennett W, Strominger J, Wiley DC (1987b) The foreign antigen binding site and T cell recognition regions of class I histocompatibility antigens. *Nature* 329 (6139): 512-518.

Bodmer WF (1987) The HLA system: structure and function. *J Clin Pathol* 40 (9): 948-958.

Borghi A, Rizzo R, Corazza M, Bertoldi AM, Bortolotti D, Sturabotti G, Virgili A, Di Luca D (2014) HLA-G 14-bp polymorphism: a possible marker of systemic treatment response in psoriasis vulgaris? Preliminary results of a retrospective study. *Dermatol Ther* 27 (5): 284-289.

Boyson JE, Erskine R, Whitman MC, Chiu M, Lau JM, Koopman LA, Valter MM, Angelisova P, Horejsi V, Strominger JL (2002) Disulfide bond-mediated dimerization of HLA-G on the cell surface. *Proc Natl Acad Sci U S A* 99 (25): 16180-16185.

Brenol CV, Veit TD, Chies JA, Xavier RM (2012) The role of the HLA-G gene and molecule on the clinical expression of rheumatologic diseases. *Rev Bras Reumatol* 52 (1): 82-91.

Briles W, McGibbon W, Irwin DM (1950) On multiple alleles affecting cellular antigens in the chicken. *Genetics* 35: 633-640.

- Briles W, McGibbon W (1948) Heterozygosity of inbred lines of chickens of two loci effecting cellular antigens. *Genetics* 33 (6): 605.
- Brown JH, Jardetzky TS, Gorga JC, Stern LJ, Urban RG, Strominger JL, Wiley DC (1993) Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. *Nature* 364 (6432): 33-39.
- Cadauid LF, Watkins DI (1997) Heirs of the jaguar and the anaconda: HLA, conquest and disease in the indigenous populations of the Americas. *Tissue Antigens* 50 (6): 702-711.
- Cao M, Yie SM, Liu J, Ye SR, Xia D, Gao E (2011) Plasma soluble HLA-G is a potential biomarker for diagnosis of colorectal, gastric, esophageal and lung cancer. *Tissue Antigens* 78 (2): 120-128.
- Cardozo DM, Vansan Marangon A, Sell AM, Laguila Visentainer JE, de Souza CA (2014) HLA and Infectious Diseases. In: Yongzhi Xi ed. *HLA and Associated Important Diseases*. InTech.
- Carreno BM, Solheim JC, Harris M, Stroynowski I, Connolly JM, Hansen TH (1995) TAP associates with a unique class I conformation, whereas calnexin associates with multiple class I forms in mouse and man. *J Immunol* 155 (10): 4726-4733.
- Casp CB, She JX, McCormack WT (2003) Genes of the LMP/TAP cluster are associated with the human autoimmune disease vitiligo. *Genes Immun* 4 (7): 492-499.
- Castelli EC, Mendes-Junior CT, Donadi EA (2007) HLA-G alleles and HLA-G 14 bp polymorphisms in a Brazilian population. *Tissue Antigens* 70:62-8
- Castelli EC, Mendes-Junior CT, Donadi EA (2007) HLA-G alleles and HLA-G 14 bp polymorphisms in a Brazilian population. *Tissue Antigens* 70 (1): 62-68.
- Castro MJ, Morales P, Rojo-Amigo R, Martinez-Laso J, Allende L, Varela P, Garcia-Berciano M, Guillen-Perales J, Arnaiz-Villena A (2000) Homozygous HLA-G\*0105N healthy individuals indicate that membrane-anchored HLA-G1 molecule is not necessary for survival. *Tissue Antigens* 56 (3): 232-239.
- Cheng CY, Su SC, Chen CH, Chen WL, Deng ST, Chung WH (2014) HLA associations and clinical implications in T-cell mediated drug hypersensitivity reactions: an updated review. *J Immunol Res* 2014: 565320.
- Chung WH, Hung SI, Hong HS, Hsieh MS, Yang LC, Ho HC, Wu JY, Chen YT (2004) Medical genetics: a marker for Stevens-Johnson syndrome. *Nature* 428 (6982): 486.
- Connolly JM, Hansen TH, Ingold AL, Potter TA (1990) Recognition by CD8 on cytotoxic T lymphocytes is ablated by several substitutions in the class I alpha 3 domain: CD8 and the T-cell receptor recognize the same class I molecule. *Proc Natl Acad Sci U S A* 87 (6): 2137-2141.

Corell A, Morales P, Varela P, Paz-Artal E, Martin-Villa JM, Martinez-Laso J, Arnaiz-Villena A (1992) Allelic diversity at the primate major histocompatibility complex DRB6 locus. *Immunogenetics* 36 (1): 33-38.

da Silva GK, Vianna P, Veit TD, Crovella S, Catamo E, Cordero EA, Mattevi VS, Lazzaretti RK, Sprinz E, Kuhmmer R et al (2014) Influence of HLA-G polymorphisms in human immunodeficiency virus infection and hepatitis C virus co-infection in Brazilian and Italian individuals. *Infect Genet Evol* 21: 418-423.

Dahl M, Perin TL, Djuriscic S, Rasmussen M, Ohlsson J, Buus S, Lindhard A, Hviid TV (2014) Soluble human leukocyte antigen-G in seminal plasma is associated with HLA-G genotype: possible implications for fertility success. *Am J Reprod Immunol* 72 (1): 89-105.

Danilovs J, Terasaki PI, Parks MS (1980) B lymphocyte isolation by thrombin-nylon wool. Los Angeles: UCLA Tissue Typing Laboratory.

Dausset J (1958) Iso-leuko-antibodies. *Acta Haematol* 20 (1-4): 156-166.

Degos L, Dausset J (1974) Human migrations and linkage disequilibrium of HLA system. *Immunogenetics* 1: 195-210.

Di Cristofaro J, Buhler S, Frassati C, Basire A, Galicher V, Baier C, et al. (2011) Linkage disequilibrium between HLA-G\*0104 and HLA-E\*0103 alleles in Tswana Pygmies. *Tissue Antigens* 77:193–200.

Donadi EA, Castelli EC, Arnaiz-Villena A, Roger M, Rey D, Moreau P (2011) Implications of the polymorphism of HLA-G on its function, regulation, evolution and disease association. *Cell Mol Life Sci* 68 (3): 369-395.

Erlich HA, Zeidler A, Chang J, Shaw S, Raffel LJ, Klitz W, Beshkov Y, Costin G, Pressman S, Bugawan T et al (1993) HLA class II alleles and susceptibility and resistance to insulin dependent diabetes mellitus in Mexican-American families. *Nat Genet* 3 (4): 358-364.

Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evol Bioinform Online* 1: 47-50.

Falk K, Rotzschke O, Stevanovic S, Jung G, Rammensee HG (1991) Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules. *Nature* 351 (6324): 290-296.

Gao GF, Willcox BE, Wyer JR, Boulter JM, O'Callaghan CA, Maenaka K, Stuart DI, Jones EY, van der Merwe PA, Bell JI et al (2000) Classical and nonclassical class I major histocompatibility complex molecules exhibit subtle conformational differences that affect binding to CD8 $\alpha$  $\alpha$ . *J Biol Chem* 275 (20): 15232-15238.

Ge YZ, Ge Q, Li MH, Shi GM, Xu X, Xu LW, Xu Z, Lu TZ, Wu R, Zhou LH et al (2014) Association between human leukocyte antigen-G 14-bp insertion/deletion polymorphism and cancer risk: a meta-analysis and systematic review. *Hum Immunol* 75 (8): 827-832.



Geraghty DE, Koller BH, Hansen JA, Orr HT (1992) The HLA class I gene family includes at least six genes and twelve pseudogenes and gene fragments. *J Immunol* 149 (6): 1934-1946.

Geraghty DE, Koller BH, Orr HT (1987) A human major histocompatibility complex class I gene that encodes a protein with a shortened cytoplasmic segment. *Proc Natl Acad Sci USA* 84 (24): 9145-9149.

Geraghty DE, Wei XH, Orr HT, Koller BH (1990) Human leukocyte antigen F (HLA-F). An expressed HLA gene composed of a class I coding sequence linked to a novel transcribed repetitive element. *J Exp Med* 171 (1): 1-18.

Gomez-Casado E, Martinez-Laso J, Arnaiz-Villena A (2006) El Sistema Principal de Histocompatibilidad humano (HLA) y el trasplante hepático. In: E.Vicente, C.Loinaz eds. *El trasplante hepático en el comienzo del milenio*. Sao Paulo: Atheneu Hispanica, pp. 37-62.

Gómez-Casado E (2002) Relaciones entre cretenses y otros pueblos mediterráneos según los genes HLA y las culturas. Tesis doctoral dirigida por Dr. A. Arnaiz-Villena y Dr. J. Martínez-Laso. Universidad Complutense de Madrid, Facultad de Ciencias Químicas.

Gomez-Casado E, Martinez-Laso J, Moscoso J, Zamora J, Martin-Villa M, Perez-Blas M, Lopez-Santalla M, Lucas GP, Silvera C, Lowy E et al (2003) Origin of Mayans according to HLA genes and the uniqueness of Amerindians. *Tissue Antigens* 61 (6): 425-436.

Gomez-Prieto P, Parga-Lozano C, Rey D, Moreno E, Arnaiz-Villena A (2010) HLA-G, -F and -E: Polymorphism, Function, and Evolution. In: Mehra NK ed. *The HLA Complex in Biology and Medicine. A Resource Book*. New Delhi: Jaypee Brothers Medical Publisehrs, pp. 159-174.

Gonzales-Galarza FF, Christmas S, Middleton D, Jones AR (2011) Allele frequency net: a database and online repository for immune gene frequencies in worldwide populations. *Nucleic Acid Research* 39: D913-D919.

Gonzalez-Galarza FF, Christmas S, Middleton D, Jones AR (2011) Allele frequency net: a database and online repository for immune gene frequencies in worldwide populations. *Nucleic Acids Res* 39 (Database issue): D913-D919.

Gonzalez-Galarza FF, Takeshita LY, Santos EJ, Kempson F, Maia MH, Silva AL, Ghattaoraya GS, Alfievic A, Jones AR, Middleton D (2015) Allele frequency net 2015 update: new features for HLA epitopes, KIR and disease and HLA adverse drug reaction associations. *Nucleic Acid Res* 39:D784-D788

Gracia D, Navarro A, Escalante JL, Tobar JA, Figuera D, Romeo CM (1996) *Trasplantes de órganos: problemas técnicos, éticos y legales*. Madrid: Publicaciones de la Universidad Pontificia Comillas.

Greenberg JH, Turner CG, Zegura SL (1986) The settlement of the Americas: a comparison of the linguistic, dental and genetic evidence. *Curr Anthropol* 27: 477-498.

HLA Nomenclature. G References (2017). [http://hla.alleles.org/data/refs/g\\_refs.html](http://hla.alleles.org/data/refs/g_refs.html). Accedido el 2-02-2017.

Horton R, Wilming L, Rand V, Lovering RC, Bruford EA, Khodiyar VK, Lush MJ, Povey S, Talbot CC, Jr., Wright MW et al (2004) Gene map of the extended human MHC. *Nat Rev Genet* 5 (12): 889-899.

Hviid TV (2006) HLA-G in human reproduction: aspects of genetics, function and pregnancy complications. *Hum Reprod Update* 12 (3): 209-232.

Hviid TV, Meldgaard M, Sorensen S, Morling N (1997) Polymorphism of exon 3 of the HLA-G gene. *Journal of Reproductive Immunology* 35(1):31-42

Ishitani A, Kishida M, Sageshima N, Yashiki S, Sonoda S, Hayami M, Smith AG, Hatake K (1999a) Re-examination of HLA-G polymorphism in African Americans. *Immunogenetics* 49:808-11

Ishitani A, Kishida M, Sageshima N, Yashiki S, Sonoda S, Hayami M, Smith AG, Hatake K (1999b) Re-examination of HLA-G polymorphism in African Americans. *Immunogenetics* 49:808-11

Ishitani A, Kishida M, Sageshima N, Yashiki S, Sonoda S, Hayami M, Smith AG, Hatake K (1999c) Re-examination of HLA-G polymorphism in African Americans. *Immunogenetics* 49:808-11

Janeway CA, Travers P, Walport M, Shlomchik MJ (2003) El sistema inmunitario en condiciones de salud y enfermedad., 2ª Edición edition. Barcelona: MASSON.

Jeong S, Park S, Park BW, Park Y, Kwon OJ, Kim HS (2014) Human leukocyte antigen-G (HLA-G) polymorphism and expression in breast cancer patients. *PLoS One* 9 (5): e98284.

Konig R, Huang LY, Germain RN (1992) MHC class II interaction with CD4 mediated by a region analogous to the MHC class I binding site for CD8. *Nature* 356 (6372): 796-798.

Koopmann JO, Hammerling GJ, Momburg F (1997) Generation, intracellular transport and loading of peptides associated with MHC class I molecules. *Curr Opin Immunol* 9 (1): 80-88.

Kumar S, Stecher G, Tamura K (2015) MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 Molecular Biology and Evolution.

Kvist S, Wiman K, Claesson L, Peterson PA, Dobberstein B (1982) Membrane insertion and oligomeric assembly of HLA-DR histocompatibility antigens. *Cell* 29 (1): 61-69.

Lajoie J, Hargrove J, Zijenah LS, Humphrey JH, Ward BJ, Roger M (2006) Genetic variants in nonclassical major histocompatibility complex class I human leukocyte antigen (HLA)-E and HLA-G molecules are associated with susceptibility to heterosexual acquisition of HIV-1. *J Infect Dis* 193 (2): 298-301.

- Lajoie J, Jeanneau A, Faucher MC, Moreau P, Roger M (2008) Characterisation of five novel HLA-G alleles with coding DNA base changes. *Tissue Antigens* 72 (5): 502-504.
- Le Discorde M, Le Danff C, Moreau P, Rouas-Freiss N, Carosella ED (2005) HLA-G\*0105N null allele encodes functional HLA-G isoforms. *Biol Reprod* 73 (2): 280-288.
- Lee YH, Bae SC, Song GG (2014) Meta-analysis of associations between functional HLA-G polymorphisms and susceptibility to systemic lupus erythematosus and rheumatoid arthritis. *Rheumatol Int*.
- Luo M, Czarnecki C, Ramdahin S, Embree J, Plummer FA (2013) HLA-G and mother-child perinatal HIV transmission. *Hum Immunol* 74 (4): 459-463.
- Lynge NL, Djuricic S, Hviid TV (2014) Controlling the Immunological Crosstalk during Conception and Pregnancy: HLA-G in Reproduction. *Front Immunol* 5: 198.
- Madden DR, Gorga JC, Strominger JL, Wiley DC (1992) The three-dimensional structure of HLA-B27 at 2.1 Å resolution suggests a general mechanism for tight peptide binding to MHC. *Cell* 70 (6): 1035-1048.
- Marks MS, Germain RN, Bonifacio JS (1995) Transient aggregation of major histocompatibility complex class II chains during assembly in normal spleen cells. *J Biol Chem* 270 (18): 10475-10481.
- Marsh SGE, Parham P, Barber LD (1999) *The HLA FactsBook*. Londres: Academic Press.
- Martinez-Laso J, Herraiz MA, Penaloza J, Barbolla ML, Jurado ML, Macedo J, Vidart J, Cervera I (2013) Promoter sequences confirm the three different evolutionary lineages described for HLA-G. *Hum Immunol* 74 (3): 383-388.
- Martinez-Laso J, Montoya F, Areces C, Moscoso J, Silvera C, Rey D, Parga-Lozano C, Gomez-Prieto P, Enriquez dS, Arnaiz-Villena A (2011) HLA in Jaidukama: an Amerindian secluded Colombian population with new haplotypes and Asian and Pacific-shared alleles. *Mol Biol Rep* 38 (6): 3689-3701.
- Martinez-Laso J, Siles N, Moscoso J, Zamora J, Serrano-Vela JJ, Ira-Cachafeiro J, Castro MJ, Serrano-Rios M, Arnaiz-Villena A (2006) Origin of Bolivian Quechua Amerindians: their relationship with other American Indians and Asians according to HLA genes. *Eur J Med Genet* 49: 169-185.
- Matte C, Lacaille J, Zijenah L, Ward B, Roger M (2000) HLA-G and HLA-E polymorphisms in an indigenous African population. The ZVITAMBO Study Group. *Hum Immunol* 61 (11): 1150-1156.
- Matte C, Lajoie J, Lacaille J, Zijenah LS, Ward BJ, Roger M (2004) Functionally active HLA-G polymorphisms are associated with the risk of heterosexual HIV-1 infection in African women. *AIDS* 18 (3): 427-431.
- Meyer D, Single R.M, Mack S.J, Lancaster A, Nelson M.P, Erlich H.A, Fernandez-Viña M, Thomson G (2007) 13th IHWS Antropology/Human Genetic Diversity Join Report.

Single locus polymorphism of classical HLA. In: John A. Hansen. ed. Immunobiology of the human MHC. Proceeding of the 13th International Histocompatibility Workshop and Conference. Seattle: IHWG PRESS, 1, pp. 653-704.

Meyer D, Thomson G (2001) How selection shapes variation of the human major histocompatibility complex: a review. *Ann Hum Genet* 65 (Pt 1): 1-26.

Mizuno S, Trapani JA, Koller BH, Dupont B, Yang SY (1988) Isolation and nucleotide sequence of a cDNA clone encoding a novel HLA class I gene. *J Immunol* 140 (11): 4024-4030.

Morales P, Corell A, Martinez-Laso J, Martin-Villa JM, Varela P, Paz-Artal E, Allende LM, Arnaiz-Villena A (1993) Three new HLA-G alleles and their linkage disequilibria with HLA-A. *Immunogenetics* 38 (5): 323-331.

Moscoso J (2006) Estudio del Complejo Principal de Histocompatibilidad en los Mayas: Relaciones con otros amerindios y con el doblamiento de America. Tesis doctoral dirigida por Dr. A. Arnaiz-Villena y Dr. J. Martinez-Laso. Universidad Complutense de Madrid. Facultad de Medicina y Hospital Universitario 12 de Octubre.

Moscoso J, Seclen S, Serrano-Vela JI, Villena A, Martinez-Laso J, Zamora J, Moreno A, Ira-Cachafeiro J, Arnaiz-Villena A (2006) HLA genes in Lamas Peruvian-Amazonian Amerindians. *Mol Immunol* 43 (11): 1881-1889.

Moussa HN, Arian SE, Sibai BM (2014) Management of hypertensive disorders in pregnancy. *Womens Health (Lond Engl)* 10 (4): 385-404.

Muro M, Alvarez-López M, Moya-Quiles M (2015) Histocompatibilidad en trasplantes. Universidad de Murcia. <https://www.um.es/biomybiotec/web/Seminarios/2008/papers/MMuro3.pdf>. Accedido el 4-12-2014.

Naji A, Menier C, Maki G, Carosella ED, Rouas-Freiss N (2012) Neoplastic B-cell growth is impaired by HLA-G/ILT2 interaction. *Leukemia* 26 (8): 1889-1892.

NCBI (2015a) HLA-G major histocompatibility complex, class I, G [ Homo sapiens (human)]. NCBI, Gene. <http://www.ncbi.nlm.nih.gov/gene/3135>. Accedido el 29-11-2014a.

NCBI (2015b) HLA-DOB major histocompatibility complex, class II, DO beta [Homo sapiens (human)]. NCBI, Gene. <http://www.ncbi.nlm.nih.gov/gene/3112>. Accedido el 8-12-2014b.

NCBI (2015c) HLA-E major histocompatibility complex, class I, E [ Homo sapiens (human)]. NCBI, Gene. <http://www.ncbi.nlm.nih.gov/gene/3133>. Accedido el 29-11-2014c.

NCBI (2015d) HLA-F major histocompatibility complex, class I, F [ Homo sapiens (human)]. NCBI, Gene. <http://www.ncbi.nlm.nih.gov/gene/3134>. Accedido el 29-11-2014d.

- Ober C, Rosinsky B, Grimsley C, van der Ven K, Robertson A, Runge A (1996) Population genetic studies of HLA-G: allele frequencies and linkage disequilibrium with HLA-A. *Journal of Reproductive Immunology* 32:111-23
- O'Callaghan CA, Tormo J, Willcox BE, Blundell CD, Jakobsen BK, Stuart DI, McMichael AJ, Bell JL, Jones EY (1998) Production, crystallization, and preliminary X-ray analysis of the human MHC class Ib molecule HLA-E. *Protein Sci* 7 (5): 1264-1266.
- Ober C, Rosinsky B, Grimsley C, van d, V, Robertson A, Runge A (1996) Population genetic studies of HLA-G: allele frequencies and linkage disequilibrium with HLA-A1. *J Reprod Immunol* 32 (2): 111-123.
- Parga-Lozano C, Rey-Medrano D, Gomez-Prieto P, Areces C, Moscoso J, Abd-El-Fatah-Khalil S, Moreno E, Arnaiz-Villena A (2011) HLA genes in Amerindian immigrants to Madrid (Spain): epidemiology and a virtual transplantation waiting list: Amerindians in Madrid (Spain). *Mol Biol Rep* 38 (4): 2263-2271.
- Park KS, Park JS, Nam JH, Bang D, Sohn S, Lee ES (2007) HLA-E\*01:01 and HLA-G\*01:01:01 reduce the risk of Behcet's disease. *Tissue Antigens* 69:139-44
- Park Y, Park Y, Kim YS, Kwon OJ, Kim HS (2012) Allele frequencies of human leukocyte antigen-G in a Korean population. *Int J Immunogenet* 39 (1): 39-45.
- Pavlos R, Mallal S, Phillips E (2012) HLA and pharmacogenetics of drug hypersensitivity. *Pharmacogenomics* 13 (11): 1285-1306.
- Payne R, Rolfs MR (1958) Fetomaternal leukocyte incompatibility. *The Journal of Clinical Investigation* 37 (12): 1756-1763.
- Perez-Luque E, Malacara JM, Olivo-Diaz A, Alaez C, Debaz H, Vazquez-Garcia M, Garay ME, Nava LE, Burguete A, Gorodezky C (2000) Contribution of HLA class II genes to end stage renal disease in mexican patients with type 2 diabetes mellitus. *Hum Immunol* 61 (10): 1031-1038.
- Petersdorf EW (2013) The major histocompatibility complex: a model for understanding graft-versus-host disease. *Blood* 122 (11): 1863-1872.
- Petzl-Erler ML, Gorodezky C, Layrisse Zeal (1997) Anthropology report for the Latin-American Region: Amerindian and admixture populations. In: Charron D ed. Paris: EDK, Genetic diversity of HLA. Funtional and medical implications., pp. 337-345.
- Pirri A, Contieri C, Benvenuti R, da Graça Bicalho M (2009) A study of HLA-G polymorphism and linkage disequilibrium in renal transplant patients and their donors. *Transplant immunology* 20:143-9
- Ploegh HL, Orr HT, Strominger JL (1981) Major histocompatibility antigens: the human (HLA-A, -B, -C) and murine (H-2K, H-2D) class I molecules. *Cell* 24 (2): 287-299.

- Radley E, Alderton RP, Kelly A, Trowsdale J, Beck S (1994) Genomic organization of HLA-DMA and HLA-DMB. Comparison of the gene organization of all six class II families in the human major histocompatibility complex. *J Biol Chem* 269 (29): 18834-18838.
- Rahimi R, Hosseini AZ, Yari F (2010) The polymorphism of human leucocyte antigen-G gene in a healthy population of Iran. *Int. J Immunogenet.* 37:269-72
- Rammensee HG, Friede T, Stevanović S (1995) MHC ligands and peptide motifs: first listing. *Immunogenetics* 41 (4): 178-228.
- Rebmann V, da Silva NF, Wagner B, Horn PA (2014) HLA-G as a tolerogenic molecule in transplantation and pregnancy. *J Immunol Res* 2014: 297073.
- Rey D, Areces C, Alonso-Rubio J, Enriquez-de-Salamanca M, Abd-El-Fatah-Khalil S, Bendikuzė N, Fernandez-Honrado M, Barbolla L, Martín-Villa JM, Arnaiz-Villena A (2013) HLA in Georgians (Caucasus) and their relationship with Eastern Mediterraneans. *Mol Biol Rep.*
- Rizzo R, Bortolotti D, Fredj NB, Rotola A, Cura F, Castellazzi M, Tamborino C, Seraceni S, Baldi E, Melchiorri L et al (2012) Role of HLA-G 14bp deletion/insertion and +3142C>G polymorphisms in the production of sHLA-G molecules in relapsing-remitting multiple sclerosis. *Hum Immunol* 73 (11): 1140-1146.
- Robinson J, Halliwell JA, Hayhurst JD, Flicek P, Parham P, Marsh SG (2015) The IPD and IMGT/HLA database: allele variant databases. *Nucleic Acids Res* 43 (Database issue): D423-D431.
- Roche PA, Marks MS, Cresswell P (1991) Formation of a nine-subunit complex by HLA class II glycoproteins and the invariant chain. *Nature* 354 (6352): 392-394.
- Rouas-Freiss N, Moreau P, LeMaoult J, Carosella ED (2014) The dual role of HLA-G in cancer. *J Immunol Res* 2014: 359748.
- Rudensky AY, Preston-Hurlburt P, Hong SC, Barlow A, Janeway CA, Jr. (1991) Sequence analysis of peptides bound to MHC class II molecules. *Nature* 353 (6345): 622-627.
- Saiki RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT, Mullis KB, Erlich HA (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 239 (4839): 487-491.
- Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular cloning: a laboratory manual*. New York: Cold Spring Harbor Laboratory Press.
- Sanchez-Velasco P, Gomez-Casado E, Martinez-Laso J, Moscoso J, Zamora J, Lowy E, Silvera C, Cemborain A, Leyva-Cobian F, Arnaiz-Villena A (2003) HLA alleles in isolated populations from North Spain: Origin of the Basques and the ancient Iberians. *Tissue Antigens* 61 (5): 384-392.



Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* 74 (12): 5463-5467.

Saper MA, Bjorkman PJ, Wiley DC (1991) Refined structure of the human histocompatibility antigen HLA-A2 at 2.6 Å resolution. *J Mol Biol* 219 (2): 277-319.

Schneider S, Roessli D, Excoffier L (2000) ARLEQUIN: a software for population genetics database analysis 2.0.

Segat L, Zupin L, Kim HY, Catamo E, Thea DM, Kankasa C, Aldrovandi GM, Kuhn L, Crovella S (2014) HLA-G 14 bp deletion/insertion polymorphism and mother-to-child transmission of HIV. *Tissue Antigens* 83 (3): 161-167.

Shao W, Yoo Y, Tang J, Zhang K, Allen S, Kaslow R (2006) Tight linkage disequilibrium between HLA-G and HLA-A alleles in native africans. *Human Immunology* 67:S118

Shiina T, Hosomichi K, Inoko H, Kulski JK (2009) The HLA genomic loci map: expression, interaction, diversity and disease. *J Hum Genet* 54 (1): 15-39.

Shiroishi M, Kuroki K, Rasubala L, Tsumoto K, Kumagai I, Kurimoto E, Kato K, Kohda D, Maenaka K (2006) Structural basis for recognition of the nonclassical MHC molecule HLA-G by the leukocyte Ig-like receptor B2 (LILRB2/LIR2/ILT4/CD85d). *Proc Natl Acad Sci U S A* 103 (44): 16412-16417.

Shiroishi M, Tsumoto K, Amano K, Shirakihara Y, Colonna M, Braud VM, Allan DS, Makadzange A, Rowland-Jones S, Willcox B et al (2003) Human inhibitory receptors Ig-like transcript 2 (ILT2) and ILT4 compete with CD8 for MHC class I binding and bind preferentially to HLA-G. *Proc Natl Acad Sci U S A* 100 (15): 8856-8861.

Shukla H, Swaroop A, Srivastava R, Weissman SM (1990) The mRNA of a human class I gene HLA G/HLA 6.0 exhibits a restricted pattern of expression. *Nucleic Acids Res* 18 (8): 2189.

Silvera C, Vargas-Alarcon G, Areces C, Rey D, Parga-Lozano C, Gomez-Prieto P, Barbolla L, Martinez-Laso J, Arnaiz-Villena A (2011) HLA genes in Wayu Amerindians from Colombia. *Immunol Invest* 40 (1): 92-100.

Sipak-Szmigiel O, Cybulski C, Wokolorczyk D, Lubinski J, Kurzawa R, Baczkowski T, Radwan M, Radwan P, Ronin-Walknowska E (2009) HLA-G polymorphism and in vitro fertilization failure in a Polish population. *Tissue Antigens* 73:348-52

Snell Gd (1948) Methods for the study of histocompatibility genes. *Journal of Genetics* 49: 87-108.

Steffensen R, Christiansen O, Bennett E, Jersild C (1998) HLA-E polymorphism in patients with recurrent spontaneous abortion. *Tissue Antigens* 52:569-72

Suarez MB, Morales P, Castro MJ, Fernandez V, Varela P, Alvarez M, Martinez-Laso J, Arnaiz-Villena A (1997) A new HLA-G allele (HLA-G\*0105N) and its distribution in the Spanish population. *Immunogenetics* 45 (6): 464-465.

Sweeney C, Kirby B (2011) Does HLA-G prevent tissue destruction in psoriasis? *Br J Dermatol* 164 (5): 1118-1119.

Teixeira AC, Mendes-Junior CT, Souza FF, Marano LA, Deghaide NH, Ferreira SC, Mente ED, Sankarankutty AK, Elias-Junior J, Castro-e-Silva et al (2013) The 14bp-deletion allele in the HLA-G gene confers susceptibility to the development of hepatocellular carcinoma in the Brazilian population. *Tissue Antigens* 81 (6): 408-413.

Tysoe-Calnon V, Grundy J, Perkins S (1991) Molecular comparisons of the B2M-binding site in Class I Major Histocompatibility Complex a-chains and proteins of related sequences. *Biochem J* 277: 359-369.

Van der Ven K, Skrablin S, Engels G, Krebs D (1998) HLA-G polymorphisms and allele frequencies in Caucasians. *Hum. Immunol.* 59:302-12

Van Ham SM, Tjin EP, Lillemeier BF, Gruneberg U, van Meijgaarden KE, Pastoors L, Verwoerd D, Tulp A, Canas B, Rahman D et al (1997) HLA-DO is a negative modulator of HLA-DM-mediated MHC class II peptide loading. *Curr Biol* 7 (12): 950-957.

Van Rood JJ, Eernisse JG, Van Leeuwen A (1958) Leucocyte antibodies in sera from pregnant women. *Nature* 181 (4625): 1735-1736.

Vargas-Alarcon G, Gamboa R, Zuniga J, Hernandez-Pacheco G, Ramos-Kuri M, Castillo E, Gomez-Casado E, Martinez-Laso J, Arnaiz-Villena A, Granados J (2000) HLA-DR4 allele frequencies on Indian and Mestizo population from Mexico. *Hum Immunol* 61 (3): 341-344.

Vargas-Alarcon G, Moscoso J, Martinez-Laso J, Rodriguez-Perez JM, Flores-Dominguez C, Serrano-Vela JJ, Moreno A, Granados J, Arnaiz-Villena A (2007) Origin of Mexican Nahuas (Aztecs) according to HLA genes and their relationships with worldwide populations. *Mol Immunol* 44 (5): 747-755.

Weber DA, Evavold BD, Jensen PE (1996) Enhanced dissociation of HLA-DR-bound peptides in the presence of HLA-DM. *Science* 274 (5287): 618-620.

Yamashita T, Fujii T, Watanabe Y, Tokunaga K, Takadoro K, Juji T, Taketani Y (1996) HLA-G gene polymorphism in a Japanese population. *Immunogenetics* 44:186-91

Yan WH, Fan LA, Yang JQ, Xu LD, Ge Y, Yao FJ (2006) HLA-G polymorphism in a Chinese Han population with recurrent spontaneous abortion. *Int. J Immunogenet.* 33:55-8

Yunis JJ, Ossa H, Salazar M, Delgado MB, Deulofeut R, de la Hoz A, Bing DH, Ramos O, Yunis EJ, Yunis EJ (1994) Major histocompatibility complex class II alleles and haplotypes and blood groups of four Amerindian tribes of northern Colombia. *Hum Immunol* 41 (4): 248-258.



# ANEXOS



**ANEXO I. CUESTIONARIO PARA DONANTES VOLUNTARIOS.**

- Consentimiento informado del donante participante en el proyecto aprobado por la comisión de investigación.

**(CARA A)**

<b>CUESTIONARIO</b>		
País y ciudad de nacimiento:		
¿Se inyecta insulina alguno de sus padres o hermanos?:	<input type="checkbox"/> Sí	<input type="checkbox"/> No
¿Ha ganado peso desde que llegó a España? (si procede)	<input type="checkbox"/> Sí	<input type="checkbox"/> No
¿Cuánto?:		
Países y ciudades donde ha vivido en América (si procede):		
¿Cuánto tiempo hace que salió de su país? (si procede):		
Año de llegada a España (si procede):		
¿Habla otra lengua además de español?:	<input type="checkbox"/> Sí	<input type="checkbox"/> No
¿Cuál?		
<b>A RELLENAR POR EL MÉDICO:</b>		
1. Perímetro cintura -	Altura -	Peso -
2. Colesterol-HDL -		
3. Tensión Arterial -		
4. Otros -		
Nombre del Facultativo –		
<input style="width: 50px; height: 30px; border: 1px solid black;" type="text"/>		

Consiento participar en este estudio epidemiológico para “Prevenir Riesgos vasculares y del corazón en la población”, en Madrid a \_\_\_\_ de \_\_ de



(CARA B)

## **ESTUDIO PARA PREVENIR RIESGOS VASCULARES Y DEL CORAZÓN EN LA POBLACIÓN**

El presente estudio es llevado por el Centro de Transfusión de Madrid, que depende de la Consejería de Sanidad de la Comunidad. Su objetivo es determinar la prevalencia de riesgo vascular y del corazón en la población de nuestra Comunidad.

Para ello se realizarán una serie de análisis y medidas antropométricas (perímetro de cintura, altura, peso) en los donantes de sangre, que así lo consientan. Todos los datos obtenidos en el presente estudio serán anónimos, sometidos a la Ley de Protección de Datos (LO 15/99 y LO 13/95) y se manejarán para obtener conclusiones poblacionales.

No obstante, si se encontrara una determinación analítica alterada, se le comunicará. **NO SUPONE NINGÚN ESFUERZO ADICIONAL A LA DONACIÓN DE SANGRE. SOLAMENTE SE LE MEDIRÁ A USTED LA CINTURA Y SE REALIZARÁ LA DETERMINACIÓN DEL COLESTEROL-HDL en una de las muestras ya extraídas con la donación de sangre.**

Le rogamos que, si desea colaborar, rellene el cuestionario y firme el consentimiento.

Firma del donante:

## ANEXO II. SIMBOLOGÍA EN EL ADN Y CÓDIGO GENÉTICO.

Bases nitrogenadas ADN	Símbolo
<i>Púricas</i>	
Guanina	G
Adenina	A
<i>Pirimidínicas</i>	
Citosina	C
Timina	T

Ambigüedades		Símbolo
2	A / C	M
	A / G	R
	A / T	W
	C / G	S
	C / T	Y
	G / T	K
3	A / C / G	V
	A / C / T	H
	A / G / T	D
	C / G / T	B
4	A / C / G / T	N

		Segunda posición					
		U	C	A	G		
Primera posición (5')	U	Phe	Ser	Tyr	Cys	U	Tercera posición (3')
						C	
		Leu		STOP	STOP	A	
					Trp	G	
	C	Leu	Pro	His	Arg	U	
						C	
				Gln		A	
						G	
	A	Ile	Thr	Asn	Ser	U	
						C	
				Lys	Arg	A	
						G	
	G	Val	Ala	Asp	Gly	U	
						C	
				Glu		A	
						G	

## ANEXO III. INFORMACIÓN SOBRE AMINOÁCIDOS

Aminoácidos	Código de tres letras	Código de una letra
Alina	Ala	A
Arginina	Arg	R
Asparagina	Asn	N
Ácido Aspártico	Asp	D
Cisteína	Cys	C
Glutamina	Gln	Q
Ácido Glutámico	Glu	E
Glicina	Gly	G
Histidina	His	H
Isoleucina	Ile	I
Leucina	Leu	L
Lisina	Lys	K
Metionina	Met	M
Fenilalanina	Phe	F
Prolina	Pro	P
Serina	Ser	S
Treonina	Thr	T
Triptófano	Trp	W
Tirosina	Tyr	Y
Valina	Val	V

## ANEXO IV. LISTADO DE ALELOS HLA-G EN LA ACTUALIDAD.

HLA-G*01:01:01:01	HLA-G*01:01:12	HLA-G*01:05N
HLA-G*01:01:01:02	HLA-G*01:01:13	HLA-G*01:06
HLA-G*01:01:01:03	HLA-G*01:01:14	HLA-G*01:07
HLA-G*01:01:01:04	HLA-G*01:01:15	HLA-G*01:08:01
HLA-G*01:01:01:05	HLA-G*01:01:16	HLA-G*01:08:02
HLA-G*01:01:01:06	HLA-G*01:01:17	HLA-G*01:09
HLA-G*01:01:02:01	HLA-G*01:01:18	HLA-G*01:10
HLA-G*01:01:02:02	HLA-G*01:01:19	HLA-G*01:11
HLA-G*01:01:03:01	HLA-G*01:01:20	HLA-G*01:12
HLA-G*01:01:03:02	HLA-G*01:01:21	HLA-G*01:13N
HLA-G*01:01:03:03	HLA-G*01:02	HLA-G*01:14
HLA-G*01:01:04	HLA-G*01:03:01:01	HLA-G*01:15
HLA-G*01:01:05	HLA-G*01:03:01:02	HLA-G*01:16
HLA-G*01:01:06	HLA-G*01:04:01	HLA-G*01:17
HLA-G*01:01:07	HLA-G*01:04:02	
HLA-G*01:01:08	HLA-G*01:04:03	
HLA-G*01:01:09	HLA-G*01:04:04	
HLA-G*01:01:11	HLA-G*01:04:05	

**ANEXO V. ABREVIATURAS MÁS FRECUENTES EN EL TEXTO.**

<b>aa</b>	Aminoácidos
<b>AINE</b>	Antiinflamatorio no esteroideo
<b>ADN</b>	Ácido desoxirribonucleico
<b>ADNc</b>	Ácido desoxirribonucleico complementario
<b>ARN</b>	Ácido ribonucleico
<b>ARNm</b>	Ácido ribonucleico mensajero
<b>ARNt</b>	Ácido ribonucleico transferente
<b>ATP</b>	Adenosín trifosfato
<b>AR</b>	Artritis reumatoide
<b>BiP</b>	Proteína de unión de la cadena pesada de la inmunoglobulina
<b>°C</b>	Grados centígrados
<b>CCT</b>	Células transicionales
<b>CHC</b>	Carcinoma hepatocelular
<b>CIITA</b>	Class II transactivator
<b>CLIP</b>	Péptido de la cadena invariante de clase II
<b>D</b>	Desequilibrio de ligamiento
<b>D'</b>	Desequilibrio de ligamiento normalizado
<b>DC</b>	Después de Cristo
<b>dNTPs</b>	Deoxinucleótidos trifosfato
<b>EICH</b>	Enfermedad de injerto contra huésped
<b>EM</b>	Esclerosis múltiple
<b>EM'</b>	Expectación maximización
<b>FH</b>	Frecuencia haplotípica
<b>HLA</b>	Antígenos leucocitarios humanos
<b>Ii</b>	Cadena invariante
<b>IL</b>	Interleuquina
<b>ISRE</b>	Interferon-stimulated response element
<b>ITIM</b>	Motivo inhibidor inmunoreceptor basado en tirosina
<b>Kd</b>	Kilodalton
<b>KIR</b>	Receptores de células NK
<b>LEIF</b>	Lupus eritematoso inducido por fármacos
<b>LES</b>	Lupus eritematoso sistémico



<b>LIE</b>	Lesión intraepitelial escamosa
<b>LILRB</b>	Receptores de leucocitos
<b>LMP</b>	Proteínas de bajo peso molecular
<b>Mb</b>	Millones de pares de bases
<b>ml</b>	Mili litro
<b>mM</b>	Mili molar
<b>MHC</b>	Complejo principal de Histocompatibilidad
<b>MIC</b>	Genes relacionados con la cadena de MHC de clase I
<b>MIIC</b>	Compartimento de MHC de clase II
<b>MPM</b>	Marcador de peso molecular
<b>ng</b>	Nano gramo
<b>NK</b>	Células <i>Natural Killer</i>
<b>p</b>	Valor estadístico de significación
<b>pb</b>	Pares de bases
<b>PCR</b>	Reacción en cadena de la Polimerasa
<b>RAF</b>	Reacciones adversas a fármacos
<b>RE</b>	Retículo endoplasmático
<b>rpm</b>	Revoluciones por minuto
<b>SIDA</b>	Síndrome de la inmunodeficiencia adquirida
<b>SNP</b>	Polimorfismo de nucleótido simple
<b>SHF</b>	Síndrome de hipersensibilidad a fármacos
<b>SSJ/NET</b>	Síndrome de Steven-Johnson / necrólisis epidérmica tóxica.
<b>TAP</b>	Transportadores asociados a la presentación antigénica
<b>TBE</b>	Tris borato EDTA
<b>TCPH</b>	Trasplante de células progenitoras hematopoyéticas
<b>TCR</b>	Receptor de la célula T
<b>TNF</b>	Factor de necrosis tumoral
<b>TM</b>	Región Transmembrana
<b>UTR</b>	Región no traducida
<b>VIH</b>	Virus de la inmunodeficiencia humana
<b>VPH</b>	Virus del papiloma humano
<b>β2-m</b>	β2-microglobulina
<b>μl</b>	Micro litro

# ANEXO VI. LISTADO DE ALELOS HLA-G EN LA POBLACIÓN DE ESTUDIO.

Muestra	Alelos HLA-G (Cuatro dígitos de resolución)		Alelos HLA-G (Seis dígitos de resolución)		País de origen
1	01:01	01:01	01:01:01	01:01:01	Argentina
2	01:04	01:04	01:04:01	01:04:04	Uruguay
3	01:01	01:04	01:01:01	01:04:04	Ecuador
4	01:01	01:01	01:01:02	01:01:02	Bolivia
5	01:01	01:03	01:01:01	01:03:01	Perú
6	01:01	01:01	01:01:01	01:01:01	Perú
7	01:01	01:06	01:01:01	01:06	Cuba
8	01:01	01:03	01:01:01	01:03:01	Paraguay
9	01:01	01:01	01:01:01	01:01:03	Perú
10	01:01	01:01	01:01:02	01:01:17	Argentina
11	01:01	01:01	01:01:02	01:01:20	Guatemala
12	01:01	01:01	01:01:01	01:01:01	Perú
13	01:01	01:01	01:01:01	01:01:01	Colombia
14	01:01	01:01	01:01:02	01:01:02	Venezuela
15	01:01	01:01	01:01:01	01:01:01	Perú
16	01:01	01:01	01:01:01	01:01:01	Colombia
17	01:01	01:01	01:01:01	01:01:01	Argentina
18	01:01	01:04	01:01:02	01:04:01	Ecuador
19	01:01	01:03	01:01:01	01:03:01	Chile
20	01:03	01:04	01:03:01	01:04:01	Ecuador
21	01:03	01:06	01:03:01	01:06	Brasil
22	01:01	01:04	01:01:08	01:04:01	Uruguay
23	01:01	01:04	01:01:02	01:04:03	Colombia
24	01:04	01:04	01:04:04	01:04:04	Ecuador
25	01:03	01:04	01:03:01	01:04:01	Ecuador
26	01:01	01:01	01:01:07	01:01:08	Méjico
27	01:01	01:01	01:01:01	01:01:02	Ecuador
28	01:01	01:06	01:01:01	01:06	Uruguay
29	01:01	01:04	01:01:02	01:04:01	Ecuador
30	01:01	01:01	01:01:01	01:01:01	Ecuador
31	01:01	01:04	01:01:02	01:04:01	Colombia
32	01:04	01:01	01:04:01	01:01:19	Rep. Dominicana
33	01:01	01:06	01:01:01	01:06	Rep. Dominicana
34	01:01	01:04	01:01:01	01:04:04	Colombia
35	01:01	01:04	01:01:03	01:04:04	Colombia
36	01:01	01:04	01:01:01	01:04:03	Colombia
37	01:01	01:04	01:01:02	01:04:01	Ecuador
38	01:01	01:06	01:01:02	01:06	Brasil
39	01:01	01:01	01:01:01	01:01:01	Méjico
40	01:01	01:01	01:01:01	01:01:01	Rep. Dominicana
41	01:01	01:01	01:01:01	01:01:01	Colombia
42	01:01	01:01	01:01:02	01:01:02	Perú
43	01:01	01:01	01:01:01	01:01:01	Argentina
44	01:01	01:05N	01:01:01	01:05N	Argentina

45	01:01	01:04	01:01:02	01:04:01	Colombia
46	01:01	01:01	01:01:02	01:01:03	Colombia
47	01:01	01:01	01:01:01	01:01:01	Colombia
48	01:01	01:04	01:01:03	01:04:01	Colombia
49	01:01	01:01	01:01:02	01:01:02	Ecuador
50	01:01	01:01	01:01:02	01:01:02	Ecuador
51	01:01	01:04	01:01:02	01:04:01	Colombia
52	01:01	01:03	01:01:03	01:03:01	Ecuador
53	01:01	01:01	01:01:02	01:01:02	Bolivia
54	01:01	01:01	01:01:01	01:01:06	Uruguay
55	01:01	01:05N	01:01:02	01:05N	Nicaragua
56	01:04	01:04	01:04:01	01:04:01	Colombia
57	01:01	01:03	01:01:02	01:03:01	Ecuador
58	01:01	01:01	01:01:01	01:01:01	Ecuador
59	01:01	01:04	01:01:06	01:04:01	Ecuador
60	01:01	01:01	01:01:01	01:01:05	Brasil
61	01:01	01:05N	01:01:08	01:05N	Colombia
62	01:06	01:06	01:06	01:06	Colombia
63	01:01	01:01	01:01:02	01:01:02	Bolivia
64	01:01	01:01	01:01:01	01:01:20	Ecuador
65	01:01	01:04	01:01:02	01:04:01	Perú
66	01:01	01:01	01:01:01	01:01:01	Perú
67	01:01	01:04	01:01:02	01:04:01	Ecuador
68	01:01	01:01	01:01:01	01:01:01	Colombia
69	01:01	01:01	01:01:02	01:01:02	Bolivia
70	01:01	01:01	01:01:02	01:01:02	Colombia
71	01:01	01:01	01:01:02	01:01:02	Ecuador
72	01:04	01:04	01:04:01	01:04:01	Colombia
73	01:04	01:01	01:04:01	01:01:02	Ecuador
74	01:01	01:03	01:01:15	01:03:01	Méjico
75	01:06	01:01	01:06	01:01:01	Perú
76	01:01	01:01	01:01:08	01:01:20	Chile
77	01:01	01:01	01:01:02	01:01:02	Perú
78	01:01	01:03	01:01:01	01:03:01	Perú
79	01:01	01:03	01:01:01	01:03:01	Colombia
80	01:01	01:01	01:01:01	01:01:01	Argentina
81	01:01	01:04	01:01:02	01:04:05	Rep. Dominicana
82	01:04	01:01	01:04:01	01:01:08	Ecuador
83	01:01	01:01	01:01:01	01:01:01	Perú
84	01:01	01:03	01:01:02	01:03:01	Venezuela
85	01:01	01:01	01:01:01	01:01:01	Colombia
86	01:01	01:04	01:01:01	01:04:03	Ecuador
87	01:01	01:06	01:01:02	01:06	Ecuador
88	01:01	01:04	01:01:08	01:01:20	Ecuador
89	01:01	01:01	01:01:01	01:01:01	Bolivia
90	01:01	01:04	01:01:02	01:04:01	Ecuador
91	01:01	01:04	01:01:02	01:04:01	Ecuador
92	01:01	01:04	01:01:08	01:04:01	Colombia
93	01:01	01:01	01:01:01	01:01:01	Perú
94	01:01	01:04	01:01:08	01:04:01	Colombia

95	01:01	01:04	01:01:01	01:04:03	Argentina
96	01:01	01:01	01:01:01	01:01:01	Colombia
97	01:01	01:04	01:01:01	01:04:04	Colombia
98	01:01	01:01	01:01:01	01:01:06	Ecuador
99	01:04	01:04	01:04:01	01:04:01	Colombia
100	01:01	01:01	01:01:02	01:01:02	Chile
101	01:04	01:06	01:04:01	01:06	Bolivia
102	01:03	01:01	01:03:01	01:01:06	Brasil
103	01:01	01:01	01:01:02	01:01:06	Ecuador
104	01:01	01:04	01:01:01	01:04:04	Colombia
105	01:01	01:03	01:01:01	01:03:01	Ecuador
106	01:01	01:01	01:01:03	01:01:03	Colombia
107	01:01	01:03	01:01:02	01:03:01	Ecuador
108	01:01	01:03	01:01:03	01:03:01	Colombia
109	01:01	01:04	01:01:02	01:04:01	Ecuador
110	01:01	01:01	01:01:02	01:01:17	Colombia
111	01:01	01:01	01:01:01	01:01:02	Colombia
112	01:01	01:01	01:01:02	01:01:02	Perú
113	01:01	01:03	01:01:01	01:03:01	Ecuador
114	01:01	01:01	01:01:01	01:01:05	Brasil
115	01:03	01:04	01:03:01	01:04:04	Bolivia
116	01:01	01:01	01:01:01	01:01:03	Perú
117	01:01	01:04	01:01:03	01:04:01	Colombia
118	01:01	01:04	-	-	Colombia
119	01:01	01:01	-	-	Ecuador
120	01:01	01:01	-	-	Ecuador
121	01:01	01:01	-	-	Ecuador
122	01:01	01:04	-	-	Honduras
123	01:01	01:04	-	-	Perú
124	01:01	01:04	-	-	Ecuador
125	01:01	01:04	-	-	Perú
126	01:01	01:01	-	-	Colombia
127	01:01	01:04	-	-	Ecuador
128	01:01	01:04	-	-	Bolivia
129	01:01	01:01	-	-	Ecuador
130	01:01	01:01	-	-	Perú
131	01:01	01:04	-	-	Colombia
132	01:01	01:01	-	-	Ecuador
133	01:01	01:01	-	-	Ecuador
134	01:01	01:01	-	-	Bolivia
135	01:01	01:04	-	-	Bolivia
136	01:01	01:01	-	-	Perú
137	01:01	01:01	-	-	Colombia
138	01:01	01:04	-	-	Rep. Dominicana
139	01:01	01:01	-	-	Chile
140	01:01	01:01	-	-	Perú
141	01:01	01:04	-	-	Ecuador
142	01:01	01:01	-	-	Colombia
143	01:01	01:01	-	-	Paraguay
144	01:01	01:01	-	-	Ecuador

145	01:01	01:04	-	-	Ecuador
146	01:01	01:01	-	-	Méjico
147	01:01	01:01	-	-	Ecuador
148	01:01	01:01	-	-	Colombia
149	01:01	01:04	-	-	Rep. Dominicana
150	01:01	01:01	-	-	Colombia
151	01:01	01:01	-	-	Ecuador
152	01:01	01:01	-	-	Colombia
153	01:01	01:01	-	-	Ecuador
154	01:01	01:01	-	-	Perú

El símbolo - indica 'no determinado'.

**ANEXO VII. CURRICULUM VITAE****Datos personales**

---

Nombre: Mercedes Enríquez de Salamanca Carrascosa.  
DNI: 53622594-A  
Fecha de Nacimiento: 31 de enero de 1986.  
Nacionalidad: España  
Teléfono: 669 417 989  
E-mail: m.enriquezdesalamanca@gmail.com

**Situación profesional actual**

---

Organismo: Universidad Complutense – Centro de Transfusión de la Comunidad de Madrid

Facultad, Escuela o Instituto: Medicina

Depto./Secc./Unidad estr.: Microbiología I (Inmunología)

Dirección postal: Avda. Complutense s/n 28040

Teléfono (indicar prefijo, número y extensión): 91 301 7356

Fax: 91 301 7229

Especialización (Códigos UNESCO): 2412.99 (Inmunogenética) 2409.03 (Genética de las Poblaciones)

Categoría profesional: Becario predoctoral

Fecha de inicio: octubre 2010

Situación administrativa: Becario predoctoral

Dedicación a tiempo completo

### Formación Académica

---

Titulación Superior: Licenciada en CC. Biológicas

Centro: Universidad Complutense de Madrid. Facultad de Ciencias Biológicas.

Fecha: 2004-2010

Postgrado: Postgrado en Ciencias Biomédicas.

Centro: Universidad Complutense de Madrid. Facultad de Medicina.

Fecha: 2010-2017

### Participación en proyectos de I+D+i

---

- Estudio epidemiológico de los genes de obesidad e histocompatibilidad y farmacogenómico HLA en Amerindios en España: genes PC-1, PPARG2, apM1, y HLA inmunosupresores y clásicos. Entidad financiadora: Fondo de Investigaciones Sanitarias (FIS), Instituto de Salud Carlos III, 2015-2017. Investigador Principal: Antonio Arnaiz Villena.
- Genes de trasplante no clásicos (HLA-E, -F, -G, -DMB y MIC), y de obesidad en Amerindios inmigrantes a España: medicina preventiva, trasplante y farmacocinética. Entidad financiadora: Fondo de Investigaciones Sanitarias (FIS), Instituto de Salud Carlos III, 2012-2014. Investigador Principal: Antonio Arnaiz Villena.
- Genes de obesidad e histocompatibilidad no clásicos (HLA-E, -F, -G) MIC y HLA-DMB en grupos Amerindios de América inmigrantes a Madrid: medicina preventiva, trasplante y farmacocinética. Entidad financiadora: Fundación Mutua Madrileña Automovilística, 2012-2014. Investigador Principal: Antonio Arnaiz Villena.

## Publicaciones o Documentos Científico-Técnicos

## A.- Artículos publicados en Revistas Internacionales:

- 1.- **Martínez-Laso J, Montoya F, Areces C, Silvera C, Rey D, Parga-Lozano C, Gomez-Prieto P, Enríquez de Salamanca M, Arnaiz-Villena A.** *HLA in Jaidukma: Amerindian secluded Colombian population with new haplotypes and Pacific-shared alleles.* Molecular Biology Reports. 38(6): 3689-3701, 2011.
- 2.- **Longas J, Martinez-Laso J, Rey D, Areces C, Gomez-Casado E, Parga-Lozano C, Luna F, Enriquez de Salamanca M, Moral P, Arnaiz-Villena A.** *Las Alpujarras region (South East Spain) HLA genes study: evidence of a probable success of 17<sup>th</sup> century repopulation of North Spain.* Molecular Biology Reports, 39(2): 1387-1394, 2012.
- 3.- **Arnaiz-Villena A, Fernández-Honrado M, Areces C, Arribas I, Coca C, Enríquez-de-Salamanca M, Parga-Lozano C, Abd-El-Fatah S, Rey D.** *Amerinidians normalized waist circumference and obesity diagnosis standarized by biochemical and HLA data.* Molecular Biology Reports, 39(4): 4875-4878, 2012.
- 4.- **Rey D, Fernandez-Honrado M, Areces C, Algora M, Abd-El-Fatah S, Enriquez-de-Salamanca M, Coca C, Arribas I, Arnaiz-Villena A.** *Ameridians show no association of PC-1 gene Gln121 allele and obesity: a thrifty gene population genetics.* Molecular Biology Reports, 39 (7):7687-7693, 2012.
- 5.- **Arnaiz-Villena A, Areces C, Rey D, Enriquez-de-Salamanca M, Alonso-Rubio J, Ruiz-del-Valle V.** *Three different North American Siskin/Goldfinch evolutionary radiations (Genus Carduelis): Pine Siskin green morphs and European Siskin in America.* The Open Ornithology Journal, 5:73-81, 2012.
- 6.- **Arnaiz-Villena A, Areces C, Enríquez-de-Salamanca M, Abd-El-Fatah Khalil S, Arribas I, Coca C, Fernández-Honrado M, Algora M, Rey D.** *Amerindians show no association of PPAR- $\gamma$ 2 gene Ala12 allele and obesity: an “unthrifty” variant population genetics.* Molecular Biology Reports, 40(2): 1767-1774, 2013.
- 7.- **Arnaiz-Villena A, Rey D, Enríquez-de-Salamanca M, Abd-El-Fatah Khalil S, Arribas I, Coca C, Fernández-Honrado M, Algora M, Areces C.** *Amerindians show association of obesity with adiponectin gene SNP45 and SNP276: population genetics of a food intake and “thrifty” gene.* Molecular Biology Reports, 40(2): 1819-1826, 2013.
- 8.- **Arnaiz-Villena A, Enríquez-de-Salamanca M, Areces C, Abd-El-Fatah Khalil S, Fernández-Honrado M, Rey D.** *HLA-G\*01:05N null allele in Mayans (Guatemala) and Uros (Titikaka Lake, Peru): evolution and population genetics.* Human Immunology, 70(4): 478-482, 2013.



- 9.- **Rey D, Parga-Lozano C, Moscoso J, Areces C, Enríquez-de-Salamanca M, Fernández-Honrado M, Abd-El-Fatah-Khalil S, Alonso-Rubio J, Arnaiz-Villena A.** *HLA genetic profile of Mapuche (Araucanian) Amerindians from Chile.* Molecular Biology Reports, 40(7): 4257-4267, 2013.
- 10.- **Rey D, Areces C, Alonso-Rubio J, Enriquez-de-Salamanca M, Abd-El-Fatah-Khalil S, Bendikuze N, Fernandez-Honrado M, Barbolla L, Martin-Villa JM, Arnaiz-Villena A.** *HLA in Georgians (Caucasus) and their relationship with Eastern Mediterraneans.* Molecular Biology Reports, 40(10):5523-5530, 2013.
- 11.- **Arnaiz-Villena A, Vargas-Alarcón G, Areces C, Enriquez-de-Salamanca M, Abd-El-Fatah-Khalil S, Fernandez-Honrado M, Marco J, Martin-Villa JM, Rey D.** *Mixtec Mexican Amerindians: an HLA alleles study for America peopling, pharmacogenomics and transplantation.* Immunological Investigations, 43(8):738-755, 2014.
- 12.- **Arnaiz-Villena A, Ruiz-del-Valle V, Gomez-Prieto P, Rey D, Enriquez-de-Salamanca M, Marco J, Muñoz E, Martin-Villa M, Areces C.** *Carduelini New Systematics: Crimson-winged Finch (Rodhopenchys sanguineus) is included in "Arid-Zone" Carduelini Finches by Mitochondrial DNA Phylogeny.* The Open Ornithology Journal, 7:55-62, 2014.
- 13.- **Arnaiz-Villena A, Areces C, ENRIQUEZ-de-Salamanca M, Abd-El-Fatah-Khalil S, Marco J, Muñoz E, Fernandez-Honrado M, Martin-Villa M, Rey D.** *Pacific Islanders and Amerindian relatedness according to HLA autosomal genes.* Internacional Journal of Modern Anthropology, 7:44-67, 2014.
- 14.- **Rey D, Amirzargar A, Areces C, ENRIQUEZ-de-Salamanca M, Marco J, Abd-El-Fatah-Khalil S, Fernandez-Honrado M, Muñoz E, Martin-Villa JM, Arnaiz-Villena A.** *Gorgan (Turkmen in Iran) HLA genetics: transplantation, pharmacogenomics and anthropology.* Immunological Investigations, 44(1):88-100, 2015.
- 15.- **Amizargar A, Rey D, Muñoz E, Palacio Grüber J, Nikbin B, Nicknam H, Khosravi F, Joshghan H, Areces C, Enríquez de Salamanca M, Martínez-Quiles N, Martin- Villa M, Arnaiz Villena A.** *Kurds HLA Genes: Its implications in Transplantation and Pharmacogenomics.* Open Medicine Journal, 2:43-47,2015.
- 16.- **Arnaiz Villena A, Enríquez de Salamanca M, Muñoz E, Vasile L, Carballo A, Martin- Villa M, Martínez-Quiles N, Areces C, Rey D.** *Easter Island and Aleuts relatensess to Amerindians and Pacific Islanders according to HLA genes.* Tissue Antigens. 85: 396. DOI: 10.1111/tan.12557.2015.
- 17.- **Arnaiz Villena A, Gomez-Prieto P, Enriquez-de-Salamanca M, Rey D, Muñoz E, Corrales A, Martin- Villa M, Martinez-Quiles N, Areces C.** *MHC variability in wild Songbirds: normal intron size and molecule evolution suggest that a minimal essential MHC does not exist in birds.* Tissue Antigens. 85: 397.DOI: 10.1111/tan.12557.2015.
- 18.- **Arnaiz Villena A, Areces C, Enriquez-de-Salamanca M, Marco J, Muñoz E, Fernandez-Honrado M, Rey D.** *Evidence of Paleohistorical Atlantic and Pacific transoceanic genetic and cultural contacts.* Human Biology. Issue supplement S60, 156:1-334. 2015.

**B.- Artículos publicados en Revistas Nacionales:**

- 1.- **Rey D, Areces C, Enriquez-de-Salamanca M, Parga-Lozano C, Abd-El-Fatah S, Fernandez M, Arnaiz-Villena A.** *Los primeros pobladores de América y sus relaciones con poblaciones del Océano Pacífico según los genes HLA.* *Inmunología*, 31:83-91, 2012.

**C.- Artículos Pendientes de Publicación en Revistas Internacionales:**

- 1.- **Rey D, Amirzargar A, Areces C, Enriquez-de-Salamanca M, Abd-El-Fatah Khalil S, Fernandez-Honrado M, Marco J, Muñiz E, Martín-Villa JM, Arnaiz-Villena A.** *Azeris HLA genetics: anthropology, transplantation and pharmacogenomics.* *Immunological Investigations*, En Prensa, 2017.

---

**Contribuciones a Congresos**

---

**Internacionales:**

- 1.- 25<sup>th</sup> European Immunogenetics and Histocompatibility Conference. Prague, Czech Republic, May 4<sup>th</sup>-7<sup>th</sup>, 2011.
- 2.- 129<sup>th</sup> Stated Meeting of the American Ornithologists' Union. Jacksonville, Florida, USA, July 24<sup>th</sup>-29<sup>th</sup>, 2011.
- 3.- 16<sup>th</sup> International HLA and Immunogenetics Workshop Report Conference, 26<sup>th</sup> European Immunogenetics and Histocompatibility Conference and 23<sup>rd</sup> British Society for Histocompatibility and Immunogenetics Conference. Liverpool, United Kingdom, May 31<sup>st</sup>- June 3<sup>rd</sup>, 2012.
- 4.- 5<sup>th</sup> North American Ornithological Conference (NAOC-V). Vancouver, Canada, August 14<sup>th</sup>-18<sup>th</sup>, 2012.
- 5.- 27<sup>th</sup> European Immunogenetics and Histocompatibility Conference. Maastricht, the Netherlands, May 11<sup>th</sup>-14<sup>th</sup>, 2013.
- 6.- 28<sup>th</sup> European Immunogenetics and Histocompatibility Conference "From genes to therapy". Stockholm, Sweden, June 25<sup>th</sup>-28<sup>th</sup>, 2014.
- 7.- 29<sup>th</sup> European Immunogenetics and Histocompatibility Conference. Geneva, Switzerland, April 26<sup>th</sup>-29<sup>th</sup>, 2015.
- 8.- American Ornithologists' Union & Cooper Ornithological Society Conference. Norman, Oklahoma. 2015.

- 9.- European Ornithologists' Union Conference. Badajoz, España. 2015.
- 10.- 84<sup>th</sup> Annual Meeting of the American Association of Physical Anthropologists. St. Louis. Missouri. 2015.
- 11.- III Iberian Congress of Biological Systematics. Vitoria-Gasteiz. España. 2015.

**Nacionales:**

- 1.- XXXVI Congreso Nacional de la Sociedad Española de Inmunología. Pamplona, 8-11 Junio 2011.
- 2.- III Congreso de la Sociedad Española de Biología Evolutiva (SESBE). Madrid, 21-25 Noviembre 2011.
- 3.- I Congreso Ibérico de Sistemática Animal. Madrid, 17-19 Enero 2012.
- 4.- XXI Congreso Español y V Ibérico de Ornitología. Vitoria, 6-9 Diciembre 2012.
- 5.- XXXVII Congreso Nacional de Inmunología. Salamanca, 29-31 Mayo 2013.
- 6.- IV Congreso de la Sociedad Española de Biología Evolutiva (SESBE). Barcelona, 27-29 Noviembre 2013.
- 7.- XXXVIII Congreso Nacional de Inmunología. Badajoz, 8-10 Mayo 2014.
- 8.- 56 Congreso de la Sociedad Española de Endocrinología y Nutrición (SEEN). Valencia, 21-23 Mayo 2014.
- 9.- XXII Congreso Español de Ornitología. Madrid, 6-9 Diciembre 2015.

**Otros méritos o aclaraciones que se desee hacer constar**

---

**Becas recibidas:**

- Beca Predoctoral de la Fundación Mutua Madrileña Automovilística con cargo al proyecto: "Genes de obesidad e histocompatibilidad no clásicos (HLA-E, -F, -G), MIC y HLA-DMB, en grupos de Amerindios de América inmigrantes a Madrid: medicina preventiva, trasplante y farmacocinética". Año 2012-2014
- Contrato con cargo a proyecto de investigación concedido por la Universidad Complutense de Madrid, Facultad de Medicina. Año 2015.

**Experiencia docente:**

- Colaboradora Honorífica del Departamento de Microbiología I (Inmunología) de la Facultad de Medicina de la Universidad Complutense de Madrid, en docencia práctica prestando servicio a diferentes asignaturas de inmunología de las licenciaturas de Biología, Farmacia, Medicina y Bioquímica (con acreditaciones). 2011-2013
- Colaboradora en tareas docentes del Departamento de Microbiología I (Inmunología) de la Facultad de Medicina de la Universidad Complutense de Madrid en docencia en seminarios, prestando servicio al máster de Inmunología (con acreditación). 2014

**Asistencia a Congresos:**

- I Congreso Ibérico de Sistemática Animal. Madrid, 17-19 Enero 2012.
- III Congreso de la Sociedad Española de Biología Evolutiva (SESBE). Madrid, 21-25 Noviembre 2011.
- XXXVI Congreso Nacional de la Sociedad Española de Inmunología. Pamplona, 8-11 Junio 2011.

## ANEXO VIII. TRABAJOS PUBLICADOS

### A- TRABAJOS PUBLICADOS CON RESULTADOS OBTENIDOS EN ESTA TESIS.

- 1- **Arnaiz-Villena A, Enríquez-de-Salamanca M, Palacio-Grüber J, Campos C, Camacho A, Martin-Villa JM, Martínez-Quiles N, Gómez Casado E, Muñiz E.** *Characterisation and functional implications of the two new HLA-G alleles found in Amerindian and Caribbean populations.* Human Immunology. 77(9):812-6. 2016. A continuación se adjunta una copia del artículo.

### B- TRABAJOS PENDIENTES DE PUBLICACIÓN CON RESULTADOS OBTENIDOS EN ESTA TESIS.

- 1- **Enríquez-de-Salamanca M, Muñiz E, Arnaiz-Villena A.** *HLA-G in Amerindians from Madrid.*



Contents lists available at ScienceDirect

journal homepage: [www.elsevier.com/locate/humimm](http://www.elsevier.com/locate/humimm)

# Characterisation and functional implications of the two new HLA-G alleles found in Amerindian and Caribbean populations

Antonio Arnaiz-Villena<sup>a,1,\*</sup>, Mercedes Enriquez-de-Salamanca<sup>a,1</sup>, Jose Palacio-Grüber<sup>a</sup>, Cristina Campos<sup>a</sup>, Alejandro Camacho<sup>a</sup>, Jose Manuel Martin-Villa<sup>b</sup>, Narcisa Martinez-Quiles<sup>b</sup>, Eduardo Gomez-Casado<sup>c</sup>, Ester Muñiz<sup>a</sup>

<sup>a</sup> Department of Immunology, University Complutense, School of Medicine, The Madrid Regional Blood Center, Madrid, Spain

<sup>b</sup> Department of Immunology, University Complutense, School of Medicine, Madrid, Spain

<sup>c</sup> Department of Inmunología Animal, Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA), Autopista A6, Hipódromo, Madrid, Spain

## ARTICLE INFO

### Article history:

Received 16 October 2015

Revised 5 January 2016

Accepted 7 January 2016

Available online xxxx

### Keywords:

Amerindians

Ecuador

Hispaniola Is

HLA-G

HLA-G and populations

HLA-G function

HLA-G immune interaction

HLA-G ligands

ILT-2

ILT-4

## ABSTRACT

HLA-G polymorphism has been found to be relatively low in all world populations. In the present paper two new HLA-G molecules are described in ancient American natives. A new HLA-G molecule from a Ecuador Amerindian individual (male) showed four codon changes with respect to HLA-G\*01:01:01. Silent changes at  $\alpha 1$  domain (residue 57, Pro, CCG  $\rightarrow$  CCA) and  $\alpha 2$  domain (residue 93, His, CAC  $\rightarrow$  CAT and residue 100, Gly, GGC  $\rightarrow$  GGT) and one productive change in  $\alpha 3$  domain (residue 219 changed from Arg to Trp). This  $\alpha 3$  change may dramatically alter HLA-G interactions with beta-2 microglobulin, CD8, ILT-2 and ILT-4 ligands present in subsets of T, B, NK, monocytes, macrophages and dendritic cells. Another HLA-G new molecule was found in a woman from Hispaniola Island, Dominican Republic (Sto Domingo): it presented a silent change at  $\alpha 2$  domain residue 107, Gly, GGA  $\rightarrow$  GGT and non-silent change at residue 178, Met  $\rightarrow$  Thr (with respect to HLA-G\*01:01:01) which is close to class I molecule/clonotypic T cell receptor interaction sites. Functional implications of these findings are discussed.

© 2016 American Society for Histocompatibility and Immunogenetics. Published by Elsevier Inc. All rights reserved.

## 1. Introduction

The First Native American are believed to have populated America coming from Asia through the Bering land bridge between 30,000 and 12,000 years before present (BP). These conclusions have been based on cultural, morphological and genetic similarities between American and Asian populations [1]. Both Siberia [1] and Mongolia [2,3] have been put forward as the most approximate places of Amerindians origin in Asia.

Greenberg et al. [4] postulated the triple migration theory for explaining the Americas peopling on linguistic bases: Amerindians (most North and South American Indians; 12,000 BP), Na-Dene (Athabascans, Navajo, Apache; 8000 BP) and Eskimo-Aleuts (6000

BP). Studies on Y chromosome showed that more than one paternal founder haplotype arrived in America during different migrations [5], probably from Siberia [6]. DNA nuclear studies have also been carried out to ascertain the origin of First Native American (Alu insertions, [7]): three identifiable clusters of people are postulated, reflecting the geographical distribution with only one wave of immigration. Also, East Asian HLA genes have been found in the Azores Island [8]; this is concordant with recent evidence showing the spread of Chinese fleet around the world in 1421 [9]. Also, the presence of South American peopling from Asia or Polynesia has been suggested because HTLV-1 virus strain shared identical sequences in Japan and in the northern coast of South America [10] and some HLA alleles may have been introduced by the same Trans-Pacific route [11]. This was also suggested by HLA genes [12–14]. Finally, both genetic [8] and archaeological [15] evidence suggests that two-way Trans-Atlantic traffic occurred before Columbus discovered America; archaeologists in New Mexico have recently found tools used 20,000 years ago in Spanish Solutrean culture [15,16]. Other archaeological common tools to America and Europe have been found [17].

\* Corresponding author at: Departamento de Inmunología, Facultad de Medicina, Universidad Complutense, Pabellón 5, planta 4. Avda. Complutense s/n, 28040 Madrid, Spain.

E-mail address: [arnaizantonio@gmail.com](mailto:arnaizantonio@gmail.com) (A. Arnaiz-Villena).

<sup>1</sup> These authors contributed equally for this work and the order of authorship is arbitrary.

<http://dx.doi.org/10.1016/j.humimm.2016.01.006>

0198-8859/© 2016 American Society for Histocompatibility and Immunogenetics. Published by Elsevier Inc. All rights reserved.

Please cite this article in press as: A. Arnaiz-Villena et al., Characterisation and functional implications of the two new HLA-G alleles found in Amerindian and Caribbean populations, Hum. Immunol. (2016), <http://dx.doi.org/10.1016/j.humimm.2016.01.006>